



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ingeniería de Sistemas e Informática
Escuela Académico Profesional de Ingeniería de Sistemas

**Implementación de un prototipo de gestión de
contenidos orientado a SEO para mejorar el
posicionamiento web de portales peruanos**

TESINA

Para optar el Título Profesional de Ingeniero de Sistemas

AUTORES

María Jesús MELÉNDEZ CUCHO

Víctor SUÁREZ BERNAOLA

ASESOR

Erwin Mac DOWALL REYNOSO

Lima, Perú

2009



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Meléndez, M. & Suárez, V. (2009). *Implementación de un prototipo de gestión de contenidos orientado a SEO para mejorar el posicionamiento web de portales peruanos*. Tesina para optar el título profesional de Ingeniero de Sistemas Escuela Académico Profesional de Ingeniería de Sistemas, Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos, Lima, Perú.

Esta tesina esta enteramente dedicado a nuestras familias. Gracias por confiar en nosotros, sin ustedes este sueño nunca hubiera sido concretado.

AGRADECIMIENTOS

Queremos agradecer a nuestros padres que han hecho posible que hayamos llegado hasta aquí y que con su motivación y paciencia, hicieran posible que podamos dedicarnos a nuestra vocación.

Por que cuando todo parecía negro nos iluminaron.

Por que nunca dejaron de estar con nosotros.

Por que nunca nos pidieron nada a cambio.

Por que son las mejores personas que conocemos.

Por eso, Gracias.

Agradecer también a nuestro orientador que con su apoyo desinteresado contribuyó para poder desarrollar éste trabajo de una manera adecuada para la comprensión y estudio de futuras generaciones.

A todas aquellas personas que indirectamente nos ayudaron para culminar este trabajo y que muchas veces constituyen un invalorable apoyo.

Y por encima de todo damos gracias a Dios.

Implementación de un Prototipo de Gestión de Contenidos orientado a SEO para mejorar el Posicionamiento Web de Portales Peruanos

RESUMEN

Este trabajo de investigación tiene como objetivo fundamental la optimización de portales Web peruanos en los Motores de Búsqueda, para que así las empresas que posean estos portales, puedan aprovecharlo mejor como un medio de estrategia de marketing en la Internet. Para lograr el objetivo mencionado se plantea la aplicación en conjunto de 2 técnicas, *Search Engine Optimization* (SEO) y *Web Content Management System* (WCMS), que en la actualidad se aplican de forma independiente, a pesar de guardar una estrecha relación, esta integración busca formar una técnica más completa para la optimización de los portales Web. En ese sentido, se realizó un análisis de las técnicas actuales de SEO, seleccionando las más adecuadas para los portales Web peruanos y además se analizó las funcionalidades básicas de un WCMS, que permita un mejor manejo de la información útil, que se requiere en el portal. Se tomó como caso de estudio, el portal Web peruano, dedicado al Desarrollo de Software, llamado Soft Brilliance S. A. C. Se analizó el posicionamiento actual del portal en los principales Motores de Búsqueda (Google, Yahoo y MSN) y el resultado de este análisis, arrojó un bajo posicionamiento como consecuencia de una aplicación deficiente de las técnicas SEO y la falta de una adecuada gestión de contenidos web. Finalmente se diseñó un prototipo fundamentado en la integración de Técnicas SEO y WCMS, que buscará optimizar el portal Web, buscando ser amigable tanto a los Motores de Búsqueda, como a los usuarios.

Palabras Claves: Posicionamiento en Motores de Búsqueda, SEO, WCMS, Motor de Búsqueda, Gestión de Contenidos Web.

Implementation of a Content Management Prototype aimed at SEO in order to improve Web Positioning of Peruvian Web Portals

ABSTRACT

The overall goal of this research project is to optimize the search engine positioning of Peruvian web portals, which allows the enterprises operating those portals, to make better use of this kind of online marketing strategy. To achieve the above mentioned objective, this paper proposes the joint application of 2 techniques, which at the moment are used independently, even though they are closely interrelated: *Search Engine Optimization* (SEO) and *Web Content Management System* (WCMS). By integrating both methods, this paper aims to develop a more complete technique for web portal optimization. In order to do so, current SEO techniques – selecting the most appropriate ones for Peruvian web portals - and the basic WCMS functionalities, which allow for a better handling of useful information required in a portal, have been analyzed. For this purpose, a case study was being conducted examining the Peruvian software development web portal Soft Brilliance S. A. C. The analysis focused, in particular, on the web portal's current positioning in the main search engines (Google, Yahoo and MSN). The results showed a low positioning for lack of SEO techniques and appropriate web content management. Consequently, a prototype based upon the integration of SEO and WCMS techniques was designed, seeking to optimize the web portal and obtaining a user-friendly result for both search engines and users.

Key Words: Search Engine Positioning, SEO, WCMS, Search Engine, Web Content Management.

ÍNDICE

ÍNDICE DE FIGURAS	X
ÍNDICE DE TABLAS.....	XV
CAPÍTULO 1: INTRODUCCIÓN.....	1
1.1 Antecedentes.....	1
1.2 Formulación del Problema.....	3
1.3 Objetivos.....	4
1.3.1 Objetivo Primario	4
1.3.2 Objetivos Secundarios	4
1.4 Justificación	5
1.5 Propuesta	6
1.6 Alcances y Limitaciones.....	7
1.7 Estructura de la Tesis.....	7
CAPÍTULO 2: MOTORES DE BÚSQUEDA (BUSCADORES).....	10
2.1 Introducción.....	10
2.2 Concepto de un Motor de Búsqueda.....	11
2.3 Sistemas de Información Documental (SID).....	12
2.4 Clasificación de los Motores de Búsqueda.....	13
2.4.1 Motores de Búsqueda propiamente dichos o Spiders	15
2.4.2 Índices Temáticos, Directorios Índices, Catálogos	16
2.4.3 Sistemas Mixtos (Motor de Búsqueda - Directorio).....	17
2.5 Anatomía de un Motor de Búsqueda	17
2.5.1 Sistema de Exploración o Rastreador	18
2.5.2 Sistema de Recuperación de Información	19
2.6 Robots, Arañas o Rastreadores	25
2.6.1 Tipos de Robots	26
2.6.2 Funcionamiento de los Robots	26
2.6.3 Indización de las páginas.....	28
2.6.4 Alineado de los Documentos (ranking) / Algoritmos de Ranking	30

CAPÍTULO 3: POSICIONAMIENTO EN BUSCADORES	34
3.1 Introducción.....	34
3.2 Posicionamiento en Buscadores	35
3.3 Posicionamiento en Motores de Búsqueda.....	36
3.4 Las fases de un Proceso de Posicionamiento en Motores de Búsqueda.....	37
3.4.1 Fase de análisis	38
3.4.2 Fase de diseño.....	39
3.4.3 Fase de implantación	39
3.4.4 Fase de seguimiento	39
3.5 Ventajas de un Buen Posicionamiento en Motores de Búsqueda.....	40
CAPÍTULO 4: TÉCNICAS EXISTENTES PARA EL POSICIONAMIENTO EN LOS MOTORES DE BÚSQUEDA	41
4.1 Introducción.....	41
4.2 Pago por Clic (PPC)	42
4.2.1 Enlaces Patrocinados	42
4.2.2 Campaña PPC.....	43
4.2.3 Funcionamiento de una Campaña PPC	44
4.2.4 Ventajas de una Campaña PPC	48
4.2.5 Desventajas de una Campaña PPC	49
4.3 Técnica SEO	50
4.3.1 Optimización para Motores de Búsqueda (Search Engine Optimization– SEO)	50
4.3.2 Criterios Básicos (Recomendaciones) para conseguir un Buen Posicionamiento	51
4.4 SEO de Sombreo Negro (<i>Black Hat SEO</i>)	62
4.5 Servicios de Empresas Privadas	63
4.6 Cuadro Comparativo entre las Técnicas Existentes	64
CAPÍTULO 5: PROPUESTA	66
5.1 Motivación.....	66
5.2 Integración de Técnicas SEO y WebCMS	68
5.2.1 Componente SEO	71
5.2.2 Componente WebCMS.....	74
5.2.3 Componente para la Gestión del Portal Web	74
5.3 Funcionalidades de la Integración Propuesta	77

5.4	Casos de Uso	79
5.5	Diagrama de Clases	80
CAPÍTULO 6: VALIDACIÓN DE LA PROPUESTA.....		81
6.1	Descripción del Prototipo	81
6.1.1	Alcance	81
6.1.2	Objetivos.....	82
6.1.3	Diseño Físico	83
6.1.4	Diagramas de despliegue	83
6.2	Funcionalidad del Prototipo	84
6.2.1	Acceso a la Administración del Contenido: Uso de Usuario y Contraseña	84
6.2.2	Menú Principal:	85
6.3	Elección del Caso de Estudio	98
6.3.1	Análisis del Estado Inicial del Caso de Estudio	99
6.3.2	Aplicando la Solución Propuesta al Caso de Estudio.....	115
6.3.3	Análisis del Resultado de la Aplicación de la Solución al Caso de Estudio	125
6.4	Resumen	148
CAPÍTULO 7: CONCLUSIONES Y TRABAJOS FUTUROS		149
7.1	Conclusiones.....	149
7.2	Trabajos Futuros	150
REFERENCIAS BIBLIOGRÁFICAS		152
ÍNDICE DE ANEXOS		155

Índice de Figuras

Figura 1.1: Porcentaje de usuarios que acceden de una forma a un portal Web.	2
Figura 1.2: Estructura de la Tesis.	9
Figura 2.1: Funcionamiento Básico de un Motor de Búsqueda.....	11
Figura 2.2: Diagrama de un Sistema de Información Documental [5].....	12
Figura 2.3: Anatomía General de un Motor de Búsqueda	18
Figura 2.4: Rastreo de los Robots de la Estructura Hipertextual en la Internet [10].....	25
Figura 2.5: Funcionamiento de los Robots dentro de la Internet [11].....	27
Figura 3.1: Origen de las visitas www.toprural.com según Google Analytic (ene/09).....	34
Figura 3.2: Resultados Naturales y Resultados Patrocinados en Google	36
Figura 4.1: Enlaces Patrocinados para una Palabra de Búsqueda en Google.....	43
Figura 4.2: Enlaces Patrocinados para una Palabra de Búsqueda en Yahoo	44
Figura 4.3: Representación del Flujo de una Subasta de Anuncios.....	45
Figura 4.4: Ofertas de Palabras Claves en una Campaña PPC.....	46
Figura 4.5: Anuncios Contextuales en Amazon.com	48
Figura 4.6: Palabra Clave “arte” – Término de Consulta en Google.....	52
Figura 4.7: Palabra Clave “arte” – Término Presente en el Contenido del Portal Arte y Cultura del Perú (http://www.arteperu.com).....	53
Figura 5.1: Aplicación de la Gestión de contenidos Web y de SEO para obtener un posicionamiento Web exitoso.....	69
Figura 5.2: Modelo Conceptual de la Integración de técnicas SEO y WebCMS	70
Figura 5.3: Uso de imágenes y no texto en el Portal Web del hotel Del Pilar Miraflores	72
Figura 5.4: Diseño del portal Web del Restaurante Hikari basado totalmente en flash	75
Figura 5.5: Registro en caché versión sólo texto.....	76
Figura 5.6: Diagrama de Casos de Uso.	79
Figura 5.7: Diagrama de Clases.....	80
Figura 6.1: Diseño Físico del prototipo.	83
Figura 6.2: Diagrama de Despliegue.	83
Figura 6.3: Página de Inicio de Sesión.	84

Figura 6.4: Menú Principal.	85
Figura 6.5: Página para la modificación de Contenido Web.	86
Figura 6.6: Subopciones para la modificación de una Página Web.	86
Figura 6.7: Configuración de propiedades de la página Web.....	87
Figura 6.8: Editor de HTML.....	88
Figura 6.9: Editor de HTML – tablas.	90
Figura 6.10: Opción de manejar enlaces en el Editor de HTML.....	91
Figura 6.11: Seleccionar la página interna o dirección externa a la que se enlazará.	91
Figura 6.12: Opciones para la gestión de imágenes.	92
Figura 6.13: Inserción de imágenes.	92
Figura 6.14: Configurar las propiedades de la imagen.	93
Figura 6.15: Editor de ‘Image Map’.	93
Figura 6.16: Opciones para la gestión de objetos multimedia.	94
Figura 6.17: Opción para la creación de páginas Web.	94
Figura 6.18: Opción para la eliminación de una página Web.....	95
Figura 6.19: Opciones del Módulo de Administración.	96
Figura 6.20: Cargar un archivo en el Portal Web.	96
Figura 6.21: Gestión del Menú Principal.....	97
Figura 6.22: Gestión del Mapa de Sitio.	97
Figura 6.23: Ayuda SEO.....	98
Figura 6.24: Buscar página principal del portal.....	100
Figura 6.25: Como se muestra en la cache de Google el Portal Soft Brilliance S.A.C.	101
Figura 6.26: Como se muestra en la cache de Google el Portal Soft Brilliance S.A.C.	102
Figura 6.27: La mala ubicación del TAG TITTLE.....	103
Figura 6.28: El texto en el TAG TITTLE se repite en la página de Inicio y en la página ‘servicios.aspx’.	104
Figura 6.29: Archivo Flash que se muestran en la página principal del portal	105
Figura 6.30: Archivo Flash que se muestran en la cabecera de las páginas del portal.....	105
Figura 6.31: La página cliente.aspx no cuenta con texto en el encabezado H1.	105
Figura 6.32: La página cliente.aspx no cuenta con texto en el encabezado H1.	106
Figura 6.33: La imagen de la página no tiene el atributo ALT.....	106

Figura 6.34: El portal Web no cuenta con el archivo robots.txt.....	107
Figura 6.35: El portal Web no cuenta con el archivo sitemap.xml.....	108
Figura 6.36: Puntos de posicionamiento del portal Web en Google.	109
Figura 6.37: El porcentaje de palabras claves por las que es posicionado en Google.	110
Figura 6.38: Punto de posicionamiento del portal Web en Yahoo.	111
Figura 6.39: El porcentaje de palabras claves por las que es posicionado en Yahoo....	111
Figura 6.40: Punto de posicionamiento del portal Web en MSN.....	112
Figura 6.41: El porcentaje de palabras claves por las que es posicionado en MSN.....	113
Figura 6.42: Puntos de posicionamiento de los tres principales Motores de Búsqueda.	114
Figura 6.43: Nuevo Diseño de la Página Principal del Portal Web de SoftBrilliance S. A. C.	116
Figura 6.44: Nuevo Diseño de las Páginas Internas del Portal Web de SoftBrilliance S. A. C.....	117
Figura 6.45: Plantilla Base orientado a SEO	118
Figura 6.46: Archivo XML para el redirect 301.....	122
Figura 6.47: Ingreso de palabras claves en las propiedades principales de la página.	123
Figura 6.48: Archivo robots.txt para el portal de SoftBrilliance S.A.C.	123
Figura 6.49: Archivo sitemap.xml para el portal de SoftBrilliance S.A.C.....	124
Figura 6.50: Búsqueda de la página principal del portal de SoftBrilliance en Google	125
Figura 6.51: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de Google	126
Figura 6.52: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de Yahoo	126
Figura 6.53: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de MSN	127
Figura 6.54: Número de páginas indexadas en Google.....	128
Figura 6.55: Número de páginas indexadas en Yahoo..	128
Figura 6.56: Número de páginas indexadas en MSN.	129
Figura 6.57: Propiedades de la Página Web “Mantenimiento.html” del Portal de SoftBrilliance S. A. C., configuradas por el sistema GECOSEO.....	130

Figura 6.58: Optimización de la página Web ‘Mantenimiento.html’ mediante los TAG TITLE y METATAG DESCRIPTION.....	131
Figura 6.59: Puntos de posicionamiento del portal Web en Google, luego de la aplicación de la Técnica WCMS y SEO.....	135
Figura 6.60: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en Google, luego de la aplicación de la Técnica WCMS y SEO.	136
Figura 6.61: Punto de posicionamiento del portal Web en Yahoo, luego de la aplicación de la Técnica WCMS y SEO.....	137
Figura 6.62: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en Yahoo, luego de la aplicación de la Técnica WCMS y SEO.	137
Figura 6.63: Punto de posicionamiento del portal Web en MSN, luego de la aplicación de la Técnica WCMS y SEO.....	138
Figura 6.64: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en MSN, luego de la aplicación de la Técnica WCMS y SEO	139
Figura 6.65: Cuenta en Google Analytics del Portal del Portal de SoftBrilliance S. A. C.....	140
Figura 6.66: Cuenta en Google Analytics del Portal del Portal de SoftBrilliance S. A. C.....	141
Figura 6.67: Reporte Estadístico del Número de Visitas Semanalmente del Portal de SoftBrilliance S. A. C.	142
Figura 6.68: Reporte Estadístico del Número de Visitas Semanalmente del Portal de SoftBrilliance S. A. C.	142
Figura 6.69: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.	143
Figura 6.70: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.	143
Figura 6.71: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.	144
Figura 6.72: Gráfico de Visitas por Ubicación Geográfica del Portal de SoftBrilliance S. A. C.	145

Figura 6.69: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.	146
---	-----

Índice de Tablas

Tabla 2.1: Características principales de Motores de Búsqueda y Directorios.	15
Tabla 2.2: Representación de un Índice Directo.....	20
Tabla 2.3: Representación de un fichero invertido	21
Tabla 2.4: Características de Rastreo de los Robots de los Principales Motores de Búsqueda [8].....	28
Tabla 2.5: Características de la Indización realizada por los Principales Motores de Búsqueda [8].....	30
Tabla 4.1: Comparación entre las Principales Técnicas Existentes.....	65
Tabla 6.1: Listado de Páginas de Inicio Estándar en los Portales Web.....	120
Tabla 6.2: Reporte de Performance de Palabras Claves en Google.	132
Tabla 6.3: Reporte de Performance de Palabras Claves en Yahoo.....	133
Tabla 6.4: Reporte de Performance de Palabras Claves en MSN.....	134

Capítulo 1: Introducción

1.1 Antecedentes

El crecimiento de las organizaciones depende en gran medida, de la búsqueda y posicionamiento en nuevos mercados. Las organizaciones concientes de ello ven la necesidad de anunciar sus productos y/o servicios en el Internet, buscando buenos resultados dentro de un entorno competitivo y globalizado.

Luego de que estas organizaciones deciden desarrollar y “colgar” su portal Web, la mayoría de ellas tienen la sensación de haber gastado dinero y no percibir mayor beneficio de éste, pues casi nadie ingresa a ellas. La situación empeora para las pequeñas empresas quienes no pueden darse el lujo de hacer una fuerte inversión y no conseguir nada.

Es así que, el poco beneficio obtenido, obliga a que los directivos se pregunten sobre la efectividad de su campaña de marketing por Internet y surjan preguntas, tales como: a) ¿Por qué la página Web no tiene muchas visitas? b) Si se construyó un portal Web atractivo al usuario, ¿por qué no llegamos a los clientes?

La respuesta radica principalmente en que los portales Web no llegan a ser vistos por los clientes potenciales, esto se debe, en gran medida, a que los involucrados en el desarrollo de los portales desconocen que implementarlo no es solamente contratar el diseño, sino también la promoción del mismo.

Por lo general, los gerentes responsables de la implementación se enfocan exclusivamente en el diseño. Pero un buen diseño no es la solución a las necesidades de marketing por Internet. Un portal Web, por sí solo, no es lo que capta a los clientes potenciales, sino que esta debe venir acompañada de una estrategia de promoción.

En ese sentido, y debido a que los usuarios, en su mayoría, acuden a los buscadores Web para localizar un determinado portal, la forma más activa de promocionar un portal Web es hacer que éste se encuentre en las primeras posiciones de las Páginas de

Resultados de los Motores de Búsqueda (SERPs, *Search Engine Results Pages*) de Google, Yahoo, Microsoft Network (MSN), Altavista. En la figura 1.1, se muestra como los usuario acceden a los portales Web.



Figura 1.1: Porcentaje de usuarios que acceden de una forma a un portal Web.

En resumen, si los usuarios usan mayormente los buscadores Web, entonces las empresas deben conseguir que estos, sepan que el portal Web existe. Además deben lograr que los buscadores consideren que el portal proporciona aquello que un usuario busca, y finalmente, deben lograr que los buscadores coloquen al portal lo antes posible entre las páginas de resultados obtenidas, es decir, las empresas deben optimizar y posicionar el portal Web en buscadores [1].

Es así que algunas empresas ya están aceptando que es necesario trabajar en una estrategia de marketing y están empezando a promocionar sus páginas Web, pero no siguen el camino adecuado. Por ejemplo, uno de los errores más comunes es cuando deciden aumentar las visitas registrando la página en los buscadores como Google y Yahoo!, lo cual sólo perjudica el posicionamiento de dicha página.

Es muy importante comprender que el registro de un portal Web en buscadores, no es lo mismo que posicionarlo en ellos. El posicionamiento en buscadores es mucho más que un simple registro y se requiere aplicar una serie de conocimientos que sólo un especialista con experiencia en este campo conoce, para que así los *spiders* (software que rastrea un sitio Web) puedan indexarlo. Por esta razón, es que el posicionamiento no es sencillo y toma varios meses alcanzar los niveles de promoción deseados.

En conclusión, una empresa puede contar con la mejor página Web del mundo, pero si no puede ser localizada a través de los buscadores, difícilmente podrá atraer el número de clientes potenciales que se espera captar.

En parte se debe a que la mayoría de diseños Web son orientados al usuario y no a los motores de búsqueda (buscadores), con lo cual se termina teniendo un portal Web visualmente llamativo, pero que probablemente ni siquiera tengan la posibilidad de aparecer en las Páginas de Resultados del Motor de Búsqueda (SERPs).

A esto, se une que las técnicas de posicionamiento en motores de búsqueda, van evolucionando rápidamente e incluyendo nuevos criterios de evaluación, que hacen del posicionamiento en Internet un trabajo cada vez más técnico y complicado.

1.2 Formulación del Problema

El posicionamiento en motores de búsqueda, se basa en la importancia de la localización. Esto se debe a que el usuario tiende a buscar un portal Web en las primeras SERPs., tanto así que aparecer luego de la tercera página, ya es considerado fuera del radio de búsqueda.

Es por eso, que se tiene la necesidad de aparecer primeros en una búsqueda, para que el usuario ubique rápidamente el portal Web que le ofrece el producto o servicio que busca. Pero si los encargados de desarrollar el portal Web no trabajan en una estrategia de marketing por Internet, será difícil que los usuarios, a los que se quiere captar, ubiquen e ingresen al mismo.

Este panorama es más notorio en los portales Web peruanos, no se trabaja en la promoción del mismo, inclusive consideran que la sola existencia de éste es suficiente para llegar a los usuarios. Además, muchas empresas no conocen las adecuadas herramientas de promoción de un portal Web, así como también, no emplean óptimamente las técnicas existentes de posicionamiento.

Por lo expuesto, el problema principal que se busca resolver con este trabajo de investigación, es que los portales Web peruanos no son mostrados en las primeras páginas de resultados en los principales motores de búsqueda, como son Google, Yahoo y MSN.

Como ya se había mencionado anteriormente, uno de los factores que contribuyen a generar este problema, es el desconocimiento o la mala aplicación de las técnicas existentes de posicionamiento, debido a que estas técnicas son utilizadas de forma independiente al momento de diseñar un portal Web y están en continuo cambio dificultando su aplicación.

1.3 Objetivos

Los objetivos de este trabajo de investigación son:

1.3.1 Objetivo Primario

Se tiene como objetivo primario, desarrollar un Prototipo de Gestión de Contenidos Web orientado a la utilización de diversas técnicas de optimización para motores de búsqueda, obteniendo como resultado que las páginas Web de un Portal Peruano aparezcan entre los mejores resultados en las búsquedas que los usuarios realicen y así ser mucho más visible para posibles clientes.

1.3.2 Objetivos Secundarios

- Facilitar la gestión y actualización de contenidos en las páginas Web en tiempo real, de manera rápida, sencilla y eficiente.
- Integrar en una sola herramienta (prototipo), las mejores técnicas y/o prácticas de Optimización para Motores de Búsqueda (SEO, *Search Engine Optimization*).

- Mejorar la forma de aplicar las técnicas existentes de SEO, de forma que los robots (*spiders*) de los motores de búsqueda; sean Google, Yahoo y MSN, sean capaces de indexar el mayor número de páginas Web y de manera correcta.
- Dar a conocer la importancia de que no basta poseer un portal Web amigable para el usuario que lo navegue, sino también ser amigable para los motores de búsqueda.

1.4 Justificación

En esta sección queremos comenzar mostrando lo que ha motivado la realización del presente trabajo de investigación.

Actualmente, los negocios tienden a emigrar total o parcialmente a la Internet. Esto se debe en gran parte a que la mayoría de personas utilizan este medio para satisfacer sus necesidades de información.

En ese sentido, los buscadores son más que solo herramientas de búsqueda a la que se puede recurrir puntualmente, son también medios para estrategias de marketing. Es por eso, que el hecho de que un portal Web no este bien posicionado en los motores de búsqueda, lleva a no poder ser encontrado por clientes potenciales. En consecuencia, no tiene sentido contar con un portal Web, si no se consigue visitas y por lo tanto es como si no estuviera promocionando sus productos y/o servicios en la Internet.

Así pues, posicionarse en las primeras posiciones de las SERPs por las palabras clave que definen los usuarios visitantes es un tema muy importante que forma parte de los pilares fundamentales para el desarrollo y promoción de un Portal Web en Internet. Por esta razón, es que se ve la necesidad de acercar a las empresas al tema de posicionamiento de una forma sencilla y así se pueda obtener un mayor provecho en comparación con sus competidores más cercanos.

Además, dentro del contexto de posicionamiento en motores de búsqueda, existen diferentes técnicas a tener en cuenta, como la utilización de urls “legibles” (Urls amigables, tanto para los usuarios visitantes como para los robots), el uso de títulos apropiados en las páginas, el etiquetado correcto de enlaces e imágenes, por mencionar

las principales. Estas técnicas están en continuo cambio y hay que investigar todavía acerca de ellas, lo cual dificulta al usuario la selección de las mejores o las más óptimas.

También, cabe mencionar que estas técnicas están plasmadas en diferentes herramientas independientes y muy numerosas; con lo cual, nos impulsa el deseo de poder desarrollar un prototipo de Gestor de Contenidos Web, con las más relevantes y mejores técnicas de optimización para motores de búsqueda y así poder posicionar en las primeras SERPs, un portal Web peruano.

1.5 Propuesta

Basándonos en el desarrollo y evolución de los métodos de identificación de contenidos (forma como exploran y recogen la información los robots de búsqueda), y el consecuente desarrollo de las técnicas de posicionamiento Web, nos hacen pensar en la necesidad de que las herramientas de gestión de contenidos para Web son capaces de permitir la aplicación conjunta de las técnicas SEO.

Por esta razón, proponemos una forma de construir portales Web optimizados para los motores de búsqueda, de tal manera que puedan lograr un buen posicionamiento en los principales buscadores.

En ese sentido, la forma de construcción propuesta, busca establecer ciertas normas del buen diseño, que permiten estructurar las páginas Web de un portal, con el fin de ser amigables tanto al usuario, así como los motores de búsqueda.

Adicionalmente a eso, permite la actualización de las páginas Web a través de algunos factores claves para el posicionamiento en motores de búsqueda, como son la elección de las palabras clave, actualización periódica del contenido, edición de los títulos de las páginas y encabezados, adición de meta datos a las páginas; con el objetivo de obtener páginas Web optimizadas.

Como validación, se pretende la implementación de un sitio Web y observar la respuesta de los motores de búsqueda a las páginas registradas. En éste se contemplan los criterios planteados en el objetivo del documento.

1.6 Alcances y Limitaciones

Este trabajo de investigación propone desarrollar un prototipo de Gestión de Contenidos Web, con las mejores practicas SEO a fin de construir páginas Web optimizadas para mejorar el Posicionamiento Web de Portales Peruanos que ofrecen productos y/o servicios.

El prototipo permite construir un portal Web amigable para los robots de los motores de búsqueda de Google, Yahoo, MSN; según el código HTML de sus páginas Web.

Además, el mayor aporte del Sistema está en orientar sus funcionalidades hacia facilitar, permitir y potenciar el posicionamiento Web, con el cual se logrará una indexación adecuada en los motores de búsqueda, un mejor ranking en las SERPs, así como garantizar la accesibilidad de los contenidos siguiendo los estándares HTML [2].

1.7 Estructura de la Tesis

Este trabajo de investigación consta de siete capítulos, tal como lo muestra la figura 1.2.

Comenzando por el presente capítulo se hace una introducción que incluye, una exposición de los principales motivos que llevaron al desarrollo de la presente tesina. También, se identifican las necesidades básicas que se debe resolver mediante el prototipo planteado de Gestión de Contenidos Web orientado a SEO. Además, se plantean los objetivos que se pretenden conseguir en el trabajo de investigación.

Los siguientes 2 capítulos sirven de base para comprender con más detalle el problema que se quiere resolver. En el Capítulo 2, se da una visión general acerca de los Motores de Búsqueda y en el Capítulo 3, se menciona los conceptos básicos de Posicionamiento Web en los principales Motores de Búsqueda.

En el capítulo 4, hace una revisión del estado del arte de las diferentes soluciones ya existentes, que pueden ser empleadas para lograr un buen posicionamiento de un portal Web.

Siguiendo, en el capítulo 5, se da a conocer la propuesta del trabajo de investigación. También, se describe el prototipo y el componente principal, SEO, que concreta la propuesta de solución, se lista las funcionalidades de este y además, se presentan todos los modelos que corresponden a la etapa de análisis, previo a la construcción del prototipo.

En el capítulo 6, se muestra, a nivel práctico, la propuesta de solución. Se presenta todos los modelos de la etapa de diseño. Además se muestra la aplicación de la propuesta sobre un caso de estudio, previamente elegido. Finalmente, mostrará un antes y un después del portal correspondiente al caso de estudio, que podrá mostrar las ventajas de la propuesta de solución, presentada en el presente trabajo de investigación.

Finalmente, el capítulo 7 recopila diferentes conclusiones en el que se detallan las principales aportaciones del trabajo desarrollado y se propone posibles líneas de investigación futuras.

Adicionalmente, se incluye una sección de Anexo donde se detallan algunos puntos de interés. En el anexo A, se hace referencia al Documento de Especificación de Casos de Uso, en el anexo B, se hace referencia a los reportes de medición de las posiciones con relación a las Palabras Claves.

Y por último, en el anexo C, se incluye los reportes de Google Analytics, que se generaron como resultado de las visitas de los usuarios para el Sitio Web de SoftBrilliance S. A. C.

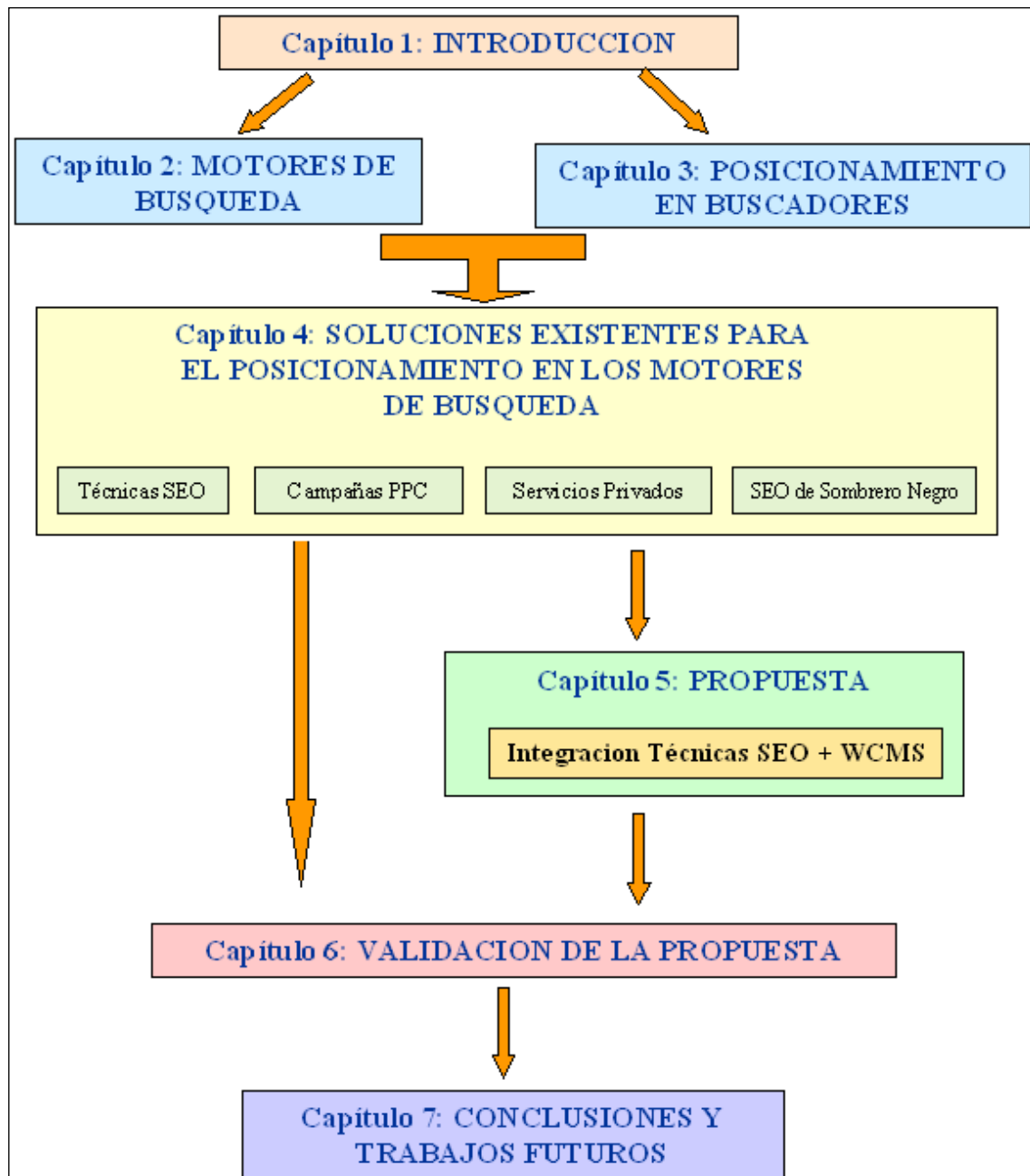


Figura 1.2: Estructura de la Tesis.

Capítulo 2: Motores de Búsqueda (Buscadores)

Esta sección nos ayuda a entender y comprender acerca de los Motores de Búsqueda y nos sirve de base para el desarrollo del presente trabajo de investigación.

2.1 Introducción

Desde hace mucho tiempo, las personas han organizado sus informaciones de tal manera que sean fáciles de encontrar y a la vez acceder de forma inmediata. La Internet, es el medio principal para publicar y colocar las informaciones a disposición de todos y los motores de búsqueda representan el medio para hacer referencia a ellas, pero a medida que va creciendo sus dimensiones resulta cada vez más difícil encontrar la información que uno va buscando y en consecuencia, más difícil para ellos catalogarla.

Por este motivo, puede parecer una paradoja, pero tener muchísima información equivale a no tener ninguna. Es así, que los motores de búsqueda han ido evolucionando y modificando sus algoritmos, que últimamente aseguran la relevancia de los resultados (Orden de Relevancia), es decir, que los primeros resultados devueltos son aquellos que probablemente busca el usuario, y que los resultados menos probables de satisfacer sus expectativas, el ruido, queden relegados a las últimas posiciones.

En ese sentido, Beatriz Aguilera, menciona en [3], que los motores de búsqueda contestan a una consulta (palabra clave de búsqueda), que el usuario ingresa en el formulario de búsqueda y devuelve un conjunto de SERPs. Estos SERPs son ordenados, en algunos casos, según la importancia que el motor de búsqueda le da a los sitios Web y es la llave del éxito de los mismos.

2.2 Concepto de un Motor de Búsqueda

También llamado “buscador Web” o simplemente “buscador”, y en inglés “*Search Engine*”; viene a ser un Sistema que gestiona un conjunto de bases de datos de Urls de diferentes contenidos en la Internet, buscando en forma booleana a través de palabras clave, es decir, recibe un pedido de búsqueda (palabras clave) y devuelve los SERPs (resultados de la búsqueda), véase la figura 2.1. Estos resultados son generalmente páginas Web, documentos, gráficos, videos, etc [4].

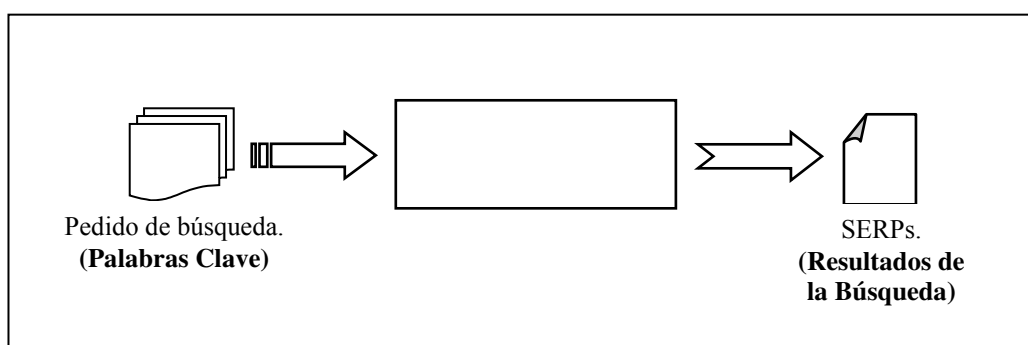


Figura 2.1: Funcionamiento Básico de un Motor de Búsqueda.

Un motor de búsqueda puede ser interno en el caso de que éste, se encuentre dentro del sistema donde se busca; o puede ser externo, en el caso de que realice la búsqueda fuera de su propio sistema. Algunos de los buscadores externos más conocidos en Internet son: Google, Yahoo, MSN, Alta Vista, Lycos, Infoseek, AllTheWeb, etc.

Una forma de entender las funciones de un motor de búsqueda consiste en concebirlo como una clase de Sistema de Información Documental (SID). En ese sentido, en la actualidad, los motores de búsqueda son los más importantes SID. Nadie podría explotar todos los servicios de la Internet, sin la ayuda de los motores de búsqueda, cuyas funciones no han dejado ni de aumentar ni de experimentar cambios desde su temprana aparición en los años 90 [5].

2.3 Sistemas de Información Documental (SID)

Una manera de comprender como funciona un motor de búsqueda es representándolo como un tipo de sistema de información documental, que se puede mostrar con el diagrama de la figura 2.2, que se muestra a continuación:

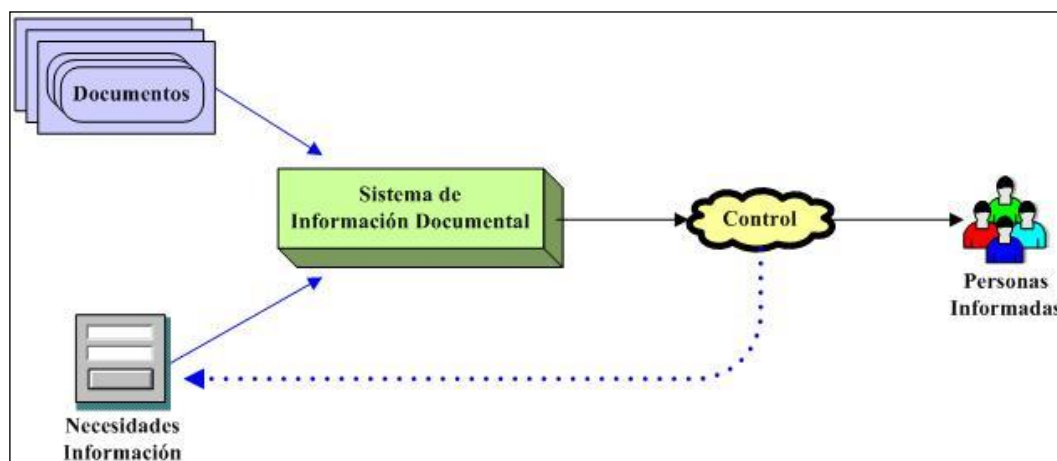


Figura 2.2: Diagrama de un Sistema de Información Documental [5].

Según el diagrama, podemos ver que un sistema documental, tal como un motor de búsqueda, necesita de dos tipos de entradas: Documentos y Necesidades de Información. Los documentos en nuestro caso son páginas Web y las necesidades de información son las preguntas de los usuarios que están buscando algo.

En cambio, la salida del Sistema es un conjunto de SERPs que, presumiblemente, contienen información capaz de satisfacer las necesidades de información del usuario, de aquí que la salida se rotule con el (optimista) texto de Personas informadas. Naturalmente, si la primera consulta no da buenos resultados, lo lógico es que el usuario realice algún intento más, usando otros términos o introduciendo algún otro tipo de modificación en la búsqueda. Esto queda representado por la nube rotulada con la palabra Control [5].

2.4 Clasificación de los Motores de Búsqueda

En [6], Baeza-Yates y Ribeiro-Neto, clasifican a los motores de búsqueda, en 3 tipos:

- Los que buscan por catálogos, como por ejemplo Yahoo.
- Los que buscan utilizando el paradigma de recuperación en texto completo, se refiere a todas las palabras de un documento. Se almacenan en un índice para su posterior recuperación; por ejemplo: Google, Altavista, Fast, Inktomi, Northern Light o Lycos.
- Y por último, los que buscan sobre otros buscadores llamados metabuscadores o multibuscadores, como por ejemplo metacrawler o metafind.

En cambio, G. Chang, en su libro *“Mining the World Wide Web”* [7], se basa en la recuperación de información por medio de palabras clave y establece 4 tipos de Motores de Búsqueda, los cuales son:

- Motores de Búsqueda, que indexan los documentos o recursos web de forma automática por medio de robots y conforman una base de datos a la que posteriormente consultará el motor cuando se realiza una consulta.
- Directorios o Índices temáticos, que son listas de recursos organizados por categorías temáticas y estructuradas de forma jerárquica que permiten visualizar los contenidos ofreciendo una lista de enlaces a las páginas que aparecen referenciadas en el buscador.
- Metabuscadores, que reúnen varios motores de búsqueda ejecutando la consulta en varios motores de forma simultánea y, por tanto, combinan los resultados de diversas fuentes, pero eliminando las referencias duplicadas y agrupando los resultados ordenándolos por pertinencia.
- Técnicas de Filtrado de Información: que es un complemento de los motores de búsqueda. El filtro permite determinar si un documento o recurso es o no relevante a priori para mantenerlo o eliminarlo de los resultados.

Y por último, Hu habla de 6 tipos de Motores de Búsqueda diferentes en [6], empleadas en la búsqueda de documentos en la Internet:

- Exploración de la Estructura Hipertextual
- Recuperación de la Información

- Metabuscadores
- Lenguajes de Consulta Basados en SQL
- Buscadores Multimedia Basados en el Contexto
- Otros

Cabe mencionar, que cada uno de los tipos de motores de búsqueda, mencionados anteriormente, presenta un diseño similar con una caja de texto en donde el usuario escribe la palabra o sentencia a buscar y un botón que ejecuta la acción de la búsqueda. Y la variación del diseño esta en función de la publicidad que oferten.

Por otro lado, en [8], se menciona que existen buscadores especializados en áreas o temas muy concretos, como son:

- Buscadores de Programas y Datos
- Buscadores de MP3
- Buscadores de Gente
- Buscadores de Noticias y Listas de Distribución
- Buscadores de Lugares
- Buscadores de Educación
- Buscadores de Software
- Buscadores de Imágenes

Pero, para la realización del presente trabajo de investigación, profundizaremos solo en 2 tipos de motores de búsqueda: motores de búsqueda propiamente dichos y directorios.

Delgado Domínguez, en [9], resume en la tabla 2.1 las características básicas de estos 2 tipos de motores de búsqueda.

	Descubrimiento de Recursos	Representación del Contenido	Representación de la Consulta	Presentación de los Resultados
Motores de Búsqueda	Principalmente de forma automática mediante robots	Indexación automática	Explícita (Mediante palabras clave o conceptos, operadores, delimitadores, etc...)	Páginas creadas de forma dinámica para consulta. Muy exhaustivos, poco precisos.
Directorios	La realizan personas	Clasificación Manual	Implícita (mediante navegación por las categorías)	Páginas creadas previamente a la consulta. Poco exhaustivos, muy precisos.

Tabla 2.1: Características principales de Motores de Búsqueda y Directorios.

2.4.1 Motores de Búsqueda propiamente dichos o Spiders

Los motores de búsqueda propiamente dichos son sistemas de búsqueda por palabras clave, que automáticamente exploran la Internet a través de “robots” que van recogiendo la información (código html) de las diferentes páginas Web y las organizan en sus bases de datos consultables por usuarios visitantes [5, 9].

Aunque hoy en día, Google se ha convertido en el motor de búsqueda más popular, la historia de éstos no empieza ni acaba con él. Ellos han ido evolucionando mostrando los resultados por Orden de Relevancia, que está determinado por diversos factores que dependen de cada motor de búsqueda. Pueden existir distintos parámetros para estos factores, para lo cual, existen muchísimas estrategias para lograr las primeras posiciones en las SERPs.

En algunos casos, los motores de búsqueda, registran todas las páginas del sitio Web, y en otros, sólo algunas. Además, las arañas van recopilando, en sus bases de datos, las actualizaciones de las últimas modificaciones de las páginas Web, cuya periodicidad depende del motor en cuestión. Por este motivo, puede suceder que algunos sitios Web no aparezcan actualizados.

La mayoría de los grandes motores de búsqueda que se conocen y se utilizan entran dentro de esta clasificación. Algunos de éstos son:

- Google.
- Altavista.
- Lycos.

2.4.2 Índices Temáticos, Directorios Índices, Catálogos

También llamados Sistemas de Navegación o Rastreadores. Son motores de búsqueda totalmente distintos a las arañas (spiders). Son sistemas de búsqueda por temas o categorías jerarquizados (aunque también suelen incluir sistemas de búsqueda por palabras clave). Se trata de bases de datos de direcciones Web elaboradas "manualmente", es decir, hay expertos que se encargan de recoger, organizar y asignar cada página Web a una categoría o tema determinado, debido a esto, el costo de operación es relativamente alto, pero su tecnología es muy barata y sencilla.

Un directorio consiste en una estructura jerárquica de clases y subclases, es decir, tenemos un campo del que derivan multitud de subcampos. Así tendríamos un campo "Música" del que podrían derivar otros subcampos como "Rock", "Clásica", "Tecno", etc. A su vez uno de estos subcampos podría seguir especializándose, dependiendo de lo general que sea. Se dice que son sistemas de navegación, ya que durante el proceso de búsqueda se va profundizando, realizando desplazamientos sucesivos entre las clases y los niveles de jerarquía. Existen por lo menos dos clases de directorios:

- Directorios de recursos.
- Directorios de directorios (metadirectorios).

En muchos casos, debido a la heterogeneidad de los temas tratados, algunas páginas presentan problemas de categorización ya que los índices debido a su base de datos de forma manual.

2.4.3 Sistemas Mixtos (Motor de Búsqueda - Directorio)

Son una mezcla entre un motor de búsqueda y un directorio. Además de tener características de buscadores, presentan las páginas registradas clasificadas en catálogos según su contenido: Informática, cultura, sociedad, que a su vez se dividen en subsecciones. Hoy en día los grandes buscadores, van en esa dirección.

En [7], se menciona esta tendencia importante, que es la de integrar la estructura de directorio con la de búsqueda. Así por ejemplo Infoseek, permite por una parte, en cada nivel de la estructura jerárquica llevar a cabo una búsqueda tanto en la categoría correspondiente como en la base de datos global generada por un robot, y por otra parte, una búsqueda en el servicio basado en el robot proporciona la opción “Related Topics” la cual conduce a las categorías correspondientes en la estructura de directorio.

En la actualidad, todos los motores tienden hacia estos sistemas mixtos

Ejemplos: Excite, Voila, Infoseek.

2.5 Anatomía de un Motor de Búsqueda

Según Lluís Codina en [5], un motor de búsqueda consiste en un conjunto de sistemas que trabajan en cooperación. Estos sistemas se pueden distribuir en dos grupos:

- Sistema de Exploración o Rastreador (en inglés: *spider* o *crawler*)
- Sistema de Recuperación de Información (en inglés: *information retrieval system*)

Y cada uno de estos sistemas se compone de otros subsistemas, tal como se muestra en la figura 2.3, que analizaremos a continuación:

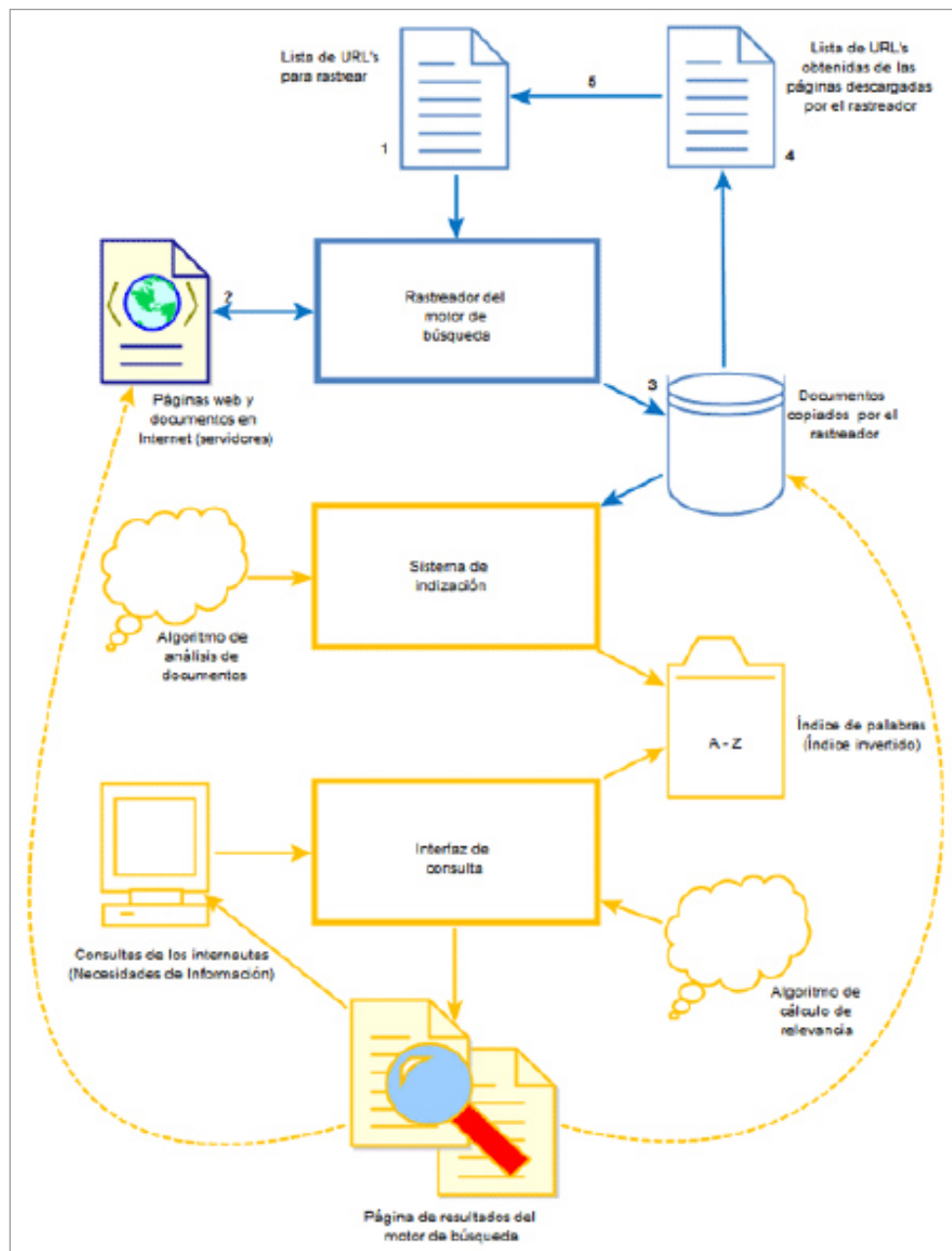


Figura 2.3: Anatomía General de un Motor de Búsqueda

2.5.1 Sistema de Exploración o Rastreador

El Rastreador, o Sistema de Exploración tiene como función de descubrir y descargar las páginas Web y otros documentos de Internet. Para ello, parte de una lista de Urls (1) (direcciones de páginas Web). A partir de ella accede a documentos (2) publicados en

Internet. Descarga esos documentos, en su mayoría páginas Web, pero también documentos ofimáticos, presentaciones, archivos de imágenes, etc. en el almacén o repositorio de documentos del Motor de Búsqueda (3). A partir de las páginas descargadas, el rastreador obtiene nuevas Urls (4) que añade a su lista de Urls pendientes de analizar (5).

Las tareas del rastreador son más complejas si se analizan con detalle, pero dado que tenemos una sección dedicada a los rastreadores, en lo que sigue, nos concentraremos en el sistema de recuperación de información.

2.5.2 Sistema de Recuperación de Información

El sistema de recuperación de información tiene a su vez dos subsistemas:

- Sistema de Indización
- Sistema de Consulta

El primero de ellos se ocupa del análisis de los documentos descargados de Internet y de la creación de los índices que permiten luego las consultas al motor; mientras que el segundo constituye la interfaz visible del MDB, precisamente la parte con la que interactúan los usuarios.

● Sistema de Indización

Si el motor de búsqueda puede responder preguntas en asombrosamente breve tiempo al que nos han acostumbrado los otros motores de búsqueda (típicamente, fracciones de segundo), es porque no exploran la Internet para los usuarios en tiempo real, es decir, en el momento de la consulta.

Un sistema que pretendiera explorar de forma secuencial los documentos de la Internet para ofrecer respuestas no tendría el más mínimo sentido. Podrían transcurrir meses o años entre la pregunta y la respuesta.

En lugar de lo anterior, lo que hacen los motores es consultar sus índices internos. Estos utilizan al menos dos clases de índices, denominados respectivamente índices directos e índices invertidos. Las consultas las resuelven los índices invertidos, mientras que los índices directos, que son los que vamos a ver ahora, se utilizan como elementos de gestión y control interno.

a. Índices Directos

En un índice directo tenemos la lista de documentos o de registros en un orden cronológico (el documento más antiguo primero) o numérico (del documento número 1 hasta el último). En la tabla 2.2, se representa esta clase de índices:

Núm. Documento	Contenido
00001	Título: Manual de lenguajes documentales Autor: Blanca Gil
00002	Título: Diseño y construcción de sitios con Dreamweaver 8 Autor: Marc Campbell
...	...
34567	Título: Aprender a razonar Autor: Fina Pizarro

Tabla 2.2: Representación de un Índice Directo

Como podemos ver, con este índice, para saber si hay un documento con las palabras “aprender” y “razonar”, tendríamos que recorrer decenas de miles de entradas del índice, para el caso de la tabla 2.2, se tendría que recorrer 34.567 registros. Lo peor es que, si el índice completo tuviera, por ejemplo, cien mil entradas, habría que recorrer las cien mil entradas del índice para saber si hay más de un documento que cumpla la condición anterior.

Como es fácil suponer, esta clase de índices no mejora mucho en relación al supuesto rastreo en tiempo real de la Internet con el que especulábamos líneas arriba.

Por tanto, un motor de búsqueda necesita complementar este índice con un índice invertido, que es el que se utiliza realmente para responder a las consultas (mientras que el índice directo se utiliza para aspectos de gestión y administración interna).

b. Índices invertidos

La estructura de un índice invertido es exactamente la inversa de la anterior (de aquí el nombre). Consiste en una lista ordenada de todas y cada una de las palabras que aparecen en los distintos documentos asociadas a los documentos concretos en los que aparecen. La estructura típica de un índice invertido, se representa en la tabla 2.3.

Término único	Frecuencia	Ubicación
Aprender	110	(34567, 01, 01)...
...
Blanca	233	(00001, 02, 01)...
...
Campbell	6	(00002, 02, 02)...
...
Razonar		(34567, 01, 03)...
...

Tabla 2.3: Representación de un fichero invertido

En donde:

- El Término único aparecen las distintas palabras de los documentos, pero solamente aparece una fila por cada palabra (aunque en el conjunto de los documentos aparezca muchas veces).
- En la columna Frecuencia tenemos el número total de veces que aparece cada término.
- Por último, en Ubicación tenemos una clave en forma de vector donde aparece el número de documento, la zona o campo donde aparece la palabra y el orden de la palabra. Habrá un vector por cada ocurrencia.

Veamos este último con más detalle. Tomemos el término “Aprender”. Tenemos que la columna de Frecuencia señala: “110”. Esto significa que “Aprender” aparece 110 veces en el conjunto de los documentos, y que habrá por tanto 110 vectores distintos en la columna Ubicación. Nosotros solamente hemos representado uno de ellos para simplificar. Tomemos ahora este mismo vector [(34567, 01, 01)] y veámoslo con detalle. El primer número, 34567, es el que identifica el documento, el segundo, 01, identifica el campo o zona del documento donde aparece la palabra, en este caso, Título.

Otros índices auxiliares ayudan al motor de búsqueda, a recorrer la lista de términos únicos con un pequeño número de operaciones de comparación. Lo mejor de todo es que esta clase de índices hace que el tiempo de respuesta sea virtualmente independiente del número de términos que aparecen en la lista.

El problema principal con estos índices es la cantidad de recursos de hardware que requieren, particularmente, si se debe dar servicio a miles de usuarios simultáneamente.

- **Sistema de Consulta - Interfaz de consulta**

Constituye la interfaz visible de los motores de búsqueda, para los usuarios que buscan información.

Las funciones que desempeña son las siguientes:

- En primer lugar, el formulario de consulta sirve para enviar la pregunta a otra parte del sistema que compara los términos de la pregunta con el índice invertido y filtra de este modo las páginas Web que contienen los términos de la pregunta.
- En segundo lugar, el Motor de Búsqueda, debe presentar los resultados de la forma más clara y eficiente posible a través de la SERP.

Por último, pero posiblemente lo más importante de todo, los resultados deben mostrarse en algún orden significativo, y de entrada podemos descartar el orden alfabético o el cronológico, dada su escasa utilidad en un contexto tan heterogéneo y con la inmensa cantidad de documentos que hay en Internet. Aquí interviene el denominado algoritmo de cálculo de relevancia, al que dedicaremos el resto de este apartado.

a. Cálculo de relevancia

En recuperación de información se considera que los distintos documentos de un fondo presentan un grado de relevancia diferente para cada pregunta o necesidad de información. Ante una pregunta dada, en un extremo tendremos a los documentos cuya relevancia será cero (por no tener ninguna relación con la pregunta). En el otro extremo tendremos (con suerte) documentos cuya relevancia será total (o casi total). En zonas intermedias tendremos documentos con diversos grados de relevancia.

Por tanto, si la expresamos en porcentaje, tendremos un espectro que irá del 0% de relevancia al 90% o al 100% de relevancia con zonas intermedias donde habrá documentos con el 10%, el 25% o el 70% de relevancia, etc.

Lo más lógico será presentar la página de resultados con los documentos ordenados según su grado de relevancia. Esta ordenación se denomina también ranking. De este modo, aunque haya miles y miles de documentos relevantes en la respuesta, podremos limitarnos a consultar los 10 o 20 primeros, puesto que se supone que son los más relevantes. Otros modos de ordenación, como el alfabético o el cronológico no nos permitirían tal economía de consulta, ya que para estar seguros que accedemos a los más relevantes ¡tendríamos que examinar la lista completa!

Los motores de búsqueda actuales suelen combinar dos grupos de criterios para determinar la relevancia de una página Web:

i. Criterios internos o intrínsecos

Los criterios internos se refieren principalmente a aspectos estadísticos o de frecuencia de ocurrencia de la palabra (o palabras) clave de la pregunta. En concreto, las páginas Web que tengan mayor densidad de la palabra clave, serán más relevantes.

Otros aspectos, como que la palabra clave aparezca rodeada de etiquetas como <h1> o <h2> también otorgan mayor importancia relativa a la página correspondiente, y sobre todo, al hecho de que la palabra clave forme parte de la URL de la página.

ii. Criterios externos o de popularidad

Los criterios externos se refieren al resultado que pueda arrojar el análisis de los enlaces de entrada de la página considerada.

Cada motor de búsqueda, tiene sus propios conjuntos de criterios y sus propias reglas para asignar pesos a cada criterio, pero en general, el número de enlaces que recibe una página suele ser uno de los más importantes, al menos a igualdad de los otros factores.

b. Página de Resultados

Por último, la respuesta del motor con la lista de documentos organizados en un ranking debe presentarse de alguna forma. De este aspecto se ocupa precisamente SERP.

Las funciones de la SERP, más allá de poder hacer clic en ellos y, por alguna razón, suelen pasar desapercibidas. Tal vez está bien que sea así, porque cada vez que se han hecho intentos en contra, parece el motor ha perdido popularidad.

En este sentido, a simple vista, la SERP ofrece una lista de páginas o documentos y ya está. Adicionalmente, en su SERP, algunos buscadores, como Google, nos muestran las siguientes informaciones y opciones:

1. Número total de páginas que contienen la palabra clave.
2. La sección de resultados de pago (enlaces patrocinados).
3. La sección de resultados no pagados (resultados orgánicos). De cada uno de estos resultados, tenemos la siguiente información:
 - Título de la página (o del documento).
 - El tipo del documento (cuando no es HTML).
 - Unas líneas de descripción del contenido del documento.
 - URL de la página.
 - Tamaño de la página Web (no de los archivos enlazados).

4. Y las siguientes opciones:

- Obtener una versión traducida de la página con traducción automática (o sea, en general muy deficiente).
- Ver la página en la caché.
- Buscar páginas con contenidos similares.
- Navegación secuencial entre los resultados o yendo a una página de resultados concreta (hasta la página 90 más o menos).
- Restringir la siguiente búsqueda a los resultados obtenidos (filtros).

2.6 Robots, Arañas o Rastreadores

Un robot es un programa que rastrea la estructura hipertextual de la Web (visitando recursivamente aquellos documentos enlazados), recogiendo información sobre las páginas que encuentra [10]. Esa información se indexa y se introduce en una base de datos que será explorada posteriormente utilizando un motor de búsqueda. A este tipo de robots se les denomina también *spiders* (arañas), rastreadores o Web crawlers. Tales robots pueden recoger más de 10 millones de páginas por día, y actualizan la información tras un período que puede ir desde unos minutos hasta 3 meses [11].

Existen diferentes tipos de robots, pero la manera general, de cómo éstos van rastreando las páginas Web, se muestra en la figura 2.4

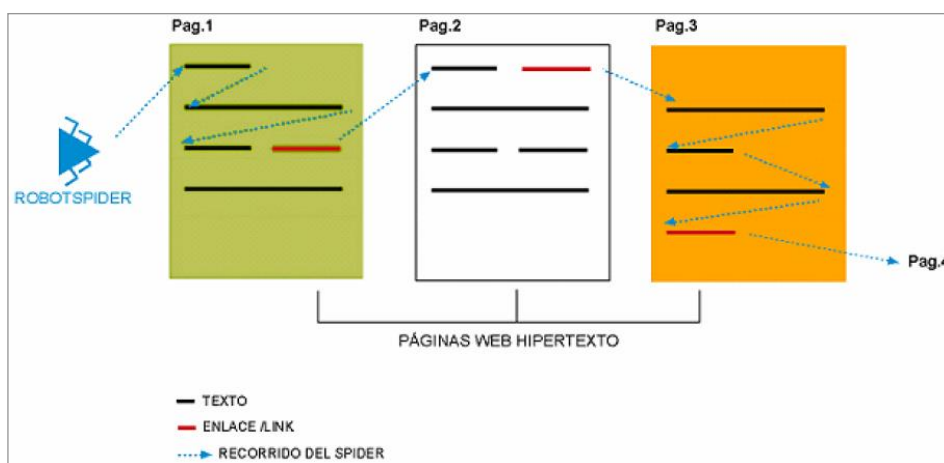


Figura 2.4: Rastreo de los Robots de la Estructura Hipertextual en la Internet [10]

2.6.1 Tipos de Robots

Si bien es cierto, el objetivo de los robots es el del descubrimiento de recursos en la Internet podemos diferenciar tipos de robots [6]:

- Knowbots: Robots programados para localizar referencias hipertexto dirigidas hacia un documento, servidor, etc., en particular. Este tipo de robots permiten evaluar el impacto de las distintas aportaciones que engrosan las distintas áreas de conocimiento presentes en la Red.
- Wanderers (vagabundos): Robots encargados de realizar análisis estadísticos, miden el crecimiento de la Red, número de servidores conectados, etc.
- Worms (gusanos): Robots encargados de la duplicación de directorios FTP para incrementar su utilidad a un número mayor de usuarios.
- WebAnts (hormigas): Conjunto de robots físicamente alejados que cooperan para la consecución de distintos objetivos, como por ejemplo para llevar a cabo una indización distribuida.

2.6.2 Funcionamiento de los Robots

Se ha comentado anteriormente que, habitualmente, el robot inicia su rastreo a partir de un conjunto de URL muy populares o enviadas explícitamente por los administradores de sitios Web, y se siguen los enlaces contenidos en esa relación inicial de páginas evitando repeticiones.

En la figura 2.5, se muestra de manera gráfica el funcionamiento general de los robots:

1. Una araña visita tu página Web, entrando por el root, lee todo el contenido y crea una lista de lo que ha encontrado.
2. La información es indexada según los algoritmos internos usados por el buscador, el cual detallaremos en la siguiente sección.
3. Esta información es llevada a una central donde se almacena.

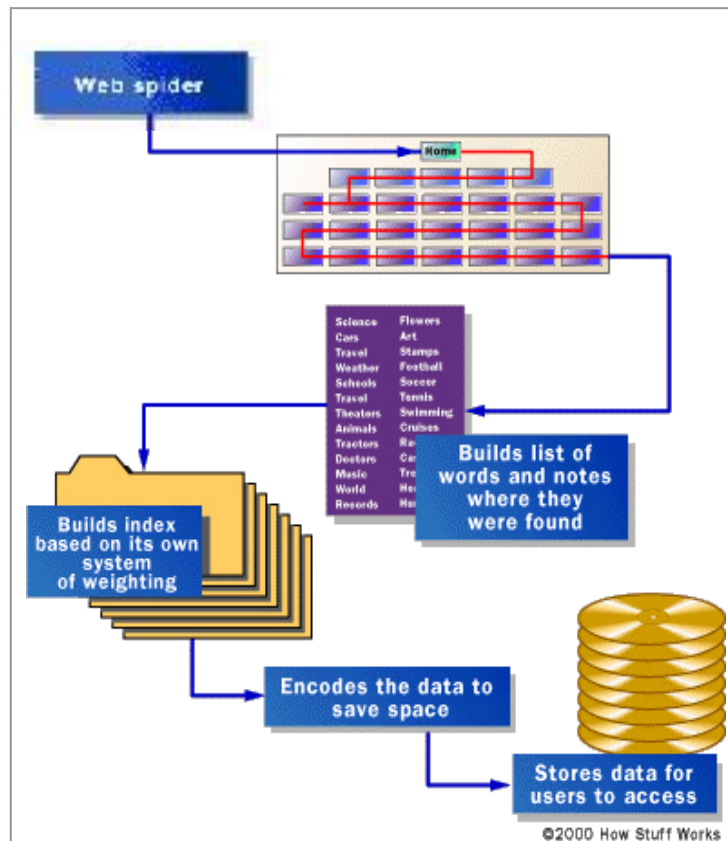


Figura 2.5: Funcionamiento de los Robots dentro de la Internet [11]

Por otro lado, citado en [6], el recorrido que realizan los robots de los motores de búsqueda, puede ser de dos modos:

- Cobertura amplia pero no profunda (*Breadth-first*).
- Cobertura vertical profunda (*Depth-first*).

A causa de la extensión de la Internet, éste genera problemas con el refresco de los índices de los motores, ya que transcurre un necesario período de tiempo entre dos análisis del mismo recurso, intervalo que varía mucho según el motor.

Analizando esta problemática, Baeza-Yates esboza, en [6], una analogía entre el índice y las estrellas del cielo: “lo que vemos en un índice jamás ha existido, ya que la luz ha viajado a lo largo de mucho tiempo hasta llegar a nuestro ojos. Cada página se indexó en un momento distinto del tiempo, pero al ir a ella obtenemos el contenido actual.

Es por ello, que algunos motores de búsqueda muestran en la respuesta, la fecha de indización de la página. Baeza-Yates, expone en [6], que alrededor del 9% de los enlaces almacenados son inválidos y Nottes cifra este porcentaje en el rango comprendido entre el 1% y 13 %, según el motor analizado. Esta cifra es objeto de estudio por varios analistas y precisa de continuas revisiones ante la naturaleza dinámica de la Internet, constituyendo uno de los criterios más significativos a la hora de ponderar la calidad de un motor de búsqueda frente a otro.

En resumen, Aguilar González, en [8], detalla en la tabla 2.4, algunas de las principales características de rastreo, por parte de los robots de los principales motores de búsqueda.

Características de Rastreo	No	Sí
Rastreo Profundo	Excite	El resto
Soporte de Marcos	Excite, FAST	El resto
Mapas de Imágenes	Excite, FAST	Alta Vista, Northern Light
Robots.txt	Ninguno	Todos
Meta datos	Excite	El resto
Rastreo por Popularidad	Ninguno	Todos
Inclusión Pagada	Excite, Google	Alta Vista, Inktomi, FAST

Tabla 2.4: Características de Rastreo de los Robots de los Principales Motores de Búsqueda [8]

2.6.3 Indización de las páginas.

A medida que los robots recopilan páginas Web, documentos, imágenes, etc; la información contenida en las mismas debe ser indizada, Delgado Domínguez opina que *“existen dos estrategias básicas, no mutuamente excluyentes, para realizar este proceso: usar información que provee el creador o editor del documento, o extraerla directamente del documento”* [9].

Y debido a que el volumen de información que gestiona un robot obliga a que el motor de búsqueda implemente algún tipo de indización automática [12]. En la práctica, los principales motores emplean ambas estrategias para disponer de una completa descripción del contenido de la página analizada.

En ese sentido, Aguilar González enumera una serie de criterios utilizados para esta descripción: *“el título del documento, los metadatos, el número de veces que se repite una palabra en un documento, algoritmos para valorar el peso del documento, etc”* [8].

La mayoría de los motores calculan el número de veces que se repiten las palabras claves en el cuerpo de una página, después escudriñan estas palabras en el nombre del dominio o en la URL, posteriormente en el título de la página, en el encabezado y en los metadatos. El orden en que se busca en cada uno de estos elementos varía en función del motor (cada uno usa sus propios algoritmos con criterios y pesos diferentes). Si el motor encuentra las palabras claves en todos estos criterios, entonces posee una razón para asignar un peso mayor al documento.

Otra metodología se basa en el número de enlaces que la misma reciba o proporcione. Aguilar González indica que la primera propuesta en esta línea es de Attardi, de la Universidad de Pisa, implementada en el motor Arianna y que ha servido de base para el desarrollo de motores que analizan los enlaces (como Google o WISEnut) [8].

Un ejemplo representativo del comportamiento de un motor clásico a la hora de indizar las páginas Web es el motor Alta Vista:

- Da prioridad alta a las palabras del título y a las palabras que están localizadas en el comienzo de la página.
- Asigna mayor peso a una palabra en un documento según su frecuencia absoluta.
- El mejor tamaño para una página está entre 4 y 8k. Considera las páginas largas como valiosas en contenido, cuando no están afectadas de “spamming”.
- Indexa las palabras claves y la descripción de los metadatos. Si no se tienen metadatos en la página, indexa las primeras 30 o 40 palabras de la página y las toma como descripción.
- Confiere una mayor prioridad a palabras ubicadas en los metadatos o a las palabras con las cuales se registran las páginas, pero no son tan relevantes como el título y el contenido.
- Es sensible a las palabras claves mayúsculas y minúsculas.
- Puede indexar un sitio que contiene marcos. Pero se debe asegurar que todas las páginas enlacen a la página principal.

Google es el mejor ejemplo de uso extensivo de los enlaces como base para mostrar los documentos a los usuarios de un motor. En este motor, la función de indización la llevan a cabo dos módulos: el indexador y el clasificador. El primero lee las páginas procedentes del storeserver⁴⁴, descomprime los documentos y selecciona los términos incluidos en los mismos.

Cada documento se convierte en un conjunto de palabras o “*hits*”, donde se graba la palabra y su posición en el documento, una aproximación de su fuente de texto y otra serie de detalles, por medio del clasificador. El indexador analiza también los enlaces incluidos en cada página Web, información necesaria para calcular el alineamiento de las páginas a la hora de la recuperación de información [13].

La Tabla 2.5 resume algunas de las principales características de la indización y los motores que las implementan.

Características de la Indización	No	Sí
Texto Completo	Ninguno	Todos
Supresión Palabras Vacías	FAST, Northern Light	Alta Vista, Excite, Inktomi, Google
Meta Descripción	Google, Northern Light	El resto
Meta Palabras Claves	Excite, FAST, Google, Northern Light	El resto
Texto Alternativo	Excite, FAST, Inktomi, Northern Light	Alta Vista, Google

Tabla 2.5: Características de la Indización realizada por los Principales Motores de Búsqueda [8]

2.6.4 Alineado de los Documentos (ranking) / Algoritmos de Ranking

El alineado constituye uno, sino el que más, de los procesos críticos a la hora de valorar la efectividad de un motor de búsqueda, ya que se trata del orden en el que el motor presenta los resultados a sus usuarios, quienes, como es lógico esperan encontrar los documentos más relevantes con sus necesidades situados entre los primeros. El motor debe ordenar el conjunto de documentos constituyente de la respuesta en función de la relevancia de estos documentos con el tema de la pregunta realizada.

En función del buen funcionamiento de su algoritmo de alineamiento (Algoritmo de Ranking), el motor será mejor o peor valorado por los usuarios del mismo. Si un motor no discrimina su respuesta en función de la relevancia con la temática objeto de la pregunta, el usuario encontrará documentos muy relevantes mezclados con otros menos relevantes e incluso con muchos nada relevantes, lo que le obligará a consultar un gran número de los documentos devueltos por el motor, teniendo que visitar muchas pantallas y perdiendo, en consecuencia, un cuantioso tiempo.

En esta situación, el usuario terminará por no recurrir a este motor de búsqueda. Si, en cambio, el motor discrimina ese grado de relación, el usuario encontrará entre los primeros documentos a los más relevantes con la temática objeto de la pregunta, por lo que aumentará su grado de satisfacción con el motor y continuará utilizándolo.

Tradicionalmente este procedimiento ha sido uno de los secretos mejor guardados por los responsables de los distintos motores de búsqueda y realmente, no se dispone de una información clara de cómo las motores lo llevan a cabo, con excepción del motor Google que ha hecho público su algoritmo PageRank [13].

Al igual que ocurría con los criterios de indización existen dos grandes grupos de algoritmos para el alineamiento, los que emplean variantes del modelo de espacio vectorial o del modelo booleano y los que siguen el principio de extensión de los enlaces [6].

Según, Baeza-Yates en [6], cita tres métodos englobados en el primer grupo, en adición al clásico esquema tf-idf: *“se denominan Booleano extendido, Vectorial extendido y Más citado. Los dos primeros son adaptaciones de los algoritmos normales de alineamiento empleados en estos modelos clásicos de recuperación de información para incluir el hecho de la existencia de enlaces entre las páginas Web. El tercero se basa únicamente en los términos incluidos en las páginas que poseen un enlace hacia las páginas de la respuesta”*.

El segundo grupo de algoritmos aporta una de las mayores diferencias conceptuales sobre el alineamiento: el uso de los enlaces de cada página (tanto los que recibe una página como los que emanan de ella). El número de enlaces que apuntan a una página sirve como una medida de su popularidad y calidad. La presencia de enlaces comunes entre un conjunto de página es también una medida de relación de los temas tratados en ellas. Dentro de esta nueva tipología de técnicas de alineamiento, identificamos tres clases:

- WebQuery: da un alineamiento a las páginas que forman la respuesta a una consulta con base en cómo de conectadas están entre ellas. Adicionalmente, extiende el conjunto de páginas de la respuesta a otra serie de páginas altamente conectadas al grupo original de respuestas.
- HITS45: alinea las páginas Web en dos tipos distintos, que guardan una relación de mutua dependencia: autoridades (páginas muy referenciadas desde otras) y hubs (o conectores, páginas desde las que se hace referencia a otras consideradas por el autor de calidad en relación a un tema).

Esta idea asume que cuando alguien establece un enlace a una página es porque la considera interesante, y que personas con intereses comunes tienden a referirse a las autoridades sobre un tema dentro de una misma página [14]. Conectores y autoridades son conceptos que se retroalimentan: mejores autoridades son inducidas por enlaces desde buenos conectores y buenos conectores vienen de enlaces desde buenas autoridades, citado en [6].

- PageRank asume que el número de enlaces que una página proporciona tiene mucho que ver con la calidad de la misma, es por ello que este algoritmo se puede resumir de la siguiente manera: “una página A tiene T_1, \dots, T_n páginas que apuntan a ella por medio de algún enlace (es decir citas). El parámetro d es un factor que se puede fijar entre 0 y 1 (generalmente se fija en 0.85). Sea $C(A)$ es número de enlaces que salen de la página A.

Entonces, el PageRank de la página A vendrá dado por la expresión:

$$PR(A) = (1-d) + d(PR(T1)/C(T1) + + PR(Tn)/C(Tn))".$$

Este cálculo puede realizarse por medio de un algoritmo iterativo y corresponde al vector propio de una matriz normalizada de enlaces en la Internet.

- PageRank está concebido como un modelo del comportamiento del usuario: si se asume que hay un "navegante aleatorio" que pasa de una página a otra sin presionar nunca el botón de “retroceder” y que, eventualmente nunca se aburriera, la probabilidad de que este navegante visitara una página determinada es precisamente su PageRank. Es decir, se trata de un modelo basado en los enlaces de las páginas y que pretende representar la forma de trabajar de los usuarios.

Otra justificación intuitiva de PageRank es que una página puede tener un alto coeficiente de PageRank si existen muchas páginas que apuntan a ella, o si hay un número algo menor de páginas que apuntan a ella pero que posean, a su vez, un alto nivel de PageRank. Lo normal es que “aquellas páginas muy citadas son páginas que vale la pena consultar y, en cambio, aquellas que sólo posean un enlace son páginas de poco interés para su consulta” [13].

Capítulo 3: Posicionamiento en Buscadores

3.1 Introducción

Cuando un usuario busca información en Internet, en especial algún producto o servicio en portales Web, y no sabe como localizarlo, lo más probable es que se acuda a un Motor de Búsqueda. Diariamente se realizan 150 millones de búsquedas en la Internet y el 71.8% del tráfico que llega a determinado portal Web lo hacen a través de los principales buscadores, vea la figura 3.1.

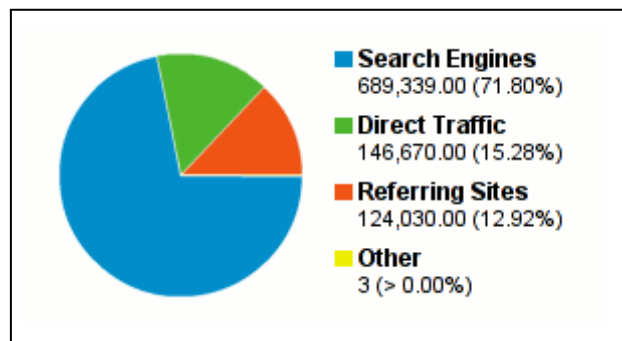


Figura 3.1: Origen de las visitas www.toprural.com según Google Analytic (ene/09)

Además, si el usuario acude a un buscador, rara vez pasan la posición 30 del resultado de la búsqueda. Esto se debe principalmente a dos motivos:

1. El usuario no dispone del tiempo necesario para visitar cientos o miles de portales Web que tratan sobre un tema determinado.
2. Las probabilidades de encontrar lo que se busca aumentan rápidamente conforme se visita portales Web del listado resultante de una búsqueda.

Los dos aspectos mencionados anteriormente, es decir: El uso de los motores de búsqueda como principal fuente de información sobre la Internet y el hecho de que los usuarios de los motores de búsqueda apenas consultan más allá de los primeros treinta o cuarenta resultados, si es que no únicamente revisa los diez primeros; son dos factores que han hecho necesario la disciplina emergente del Posicionamiento en Buscadores, que no es otra cosa que ubicar información en los SERPs. [5].

Los factores anteriores ha generado la legítima preocupación, por parte de los responsables del portal, para aparecer bien posicionados en los motores de búsqueda (y directorios, aunque en menor medida).

3.2 Posicionamiento en Buscadores

El posicionamiento en buscadores es un proceso dinámico que presenta una serie de fases. Al mismo tiempo, no existe en realidad un final en el posicionamiento de un Portal Web ya que es necesario monitorizar con cierta frecuencia el éxito de las acciones de posicionamiento y realizar acciones de corrección si es necesario.

Algunos buscadores generan dos tipos distintos de resultados que aparecen claramente diferenciadas en dos zonas, la primera zona es de resultados obtenidos en función de las palabras clave introducidas en la búsqueda y ordenados de acuerdo con el algoritmo de ordenación correspondiente a cada buscador. Este algoritmo depende de muchos factores, pero ninguno de ellos es el pago realizado por las empresas que aparecen en el listado. Estos resultados son llamados orgánicos o naturales.

En el listado de resultados aparece también el listado de anuncios bajo el título de “enlaces patrocinados”. Están situados en una zona bien diferenciada y también son activados en función de las palabras clave usadas.

Los buscadores Google y Yahoo diferencian claramente las dos zonas de resultados y aunque se indique de forma clara que se trata de enlaces patrocinados, muchos usuarios no identifican que se trate de publicidad. La zona de enlaces patrocinados en ambos buscadores es idéntica en forma de L invertida y con la frase “Enlaces patrocinados”, tal como se ve en la figura 3.2.

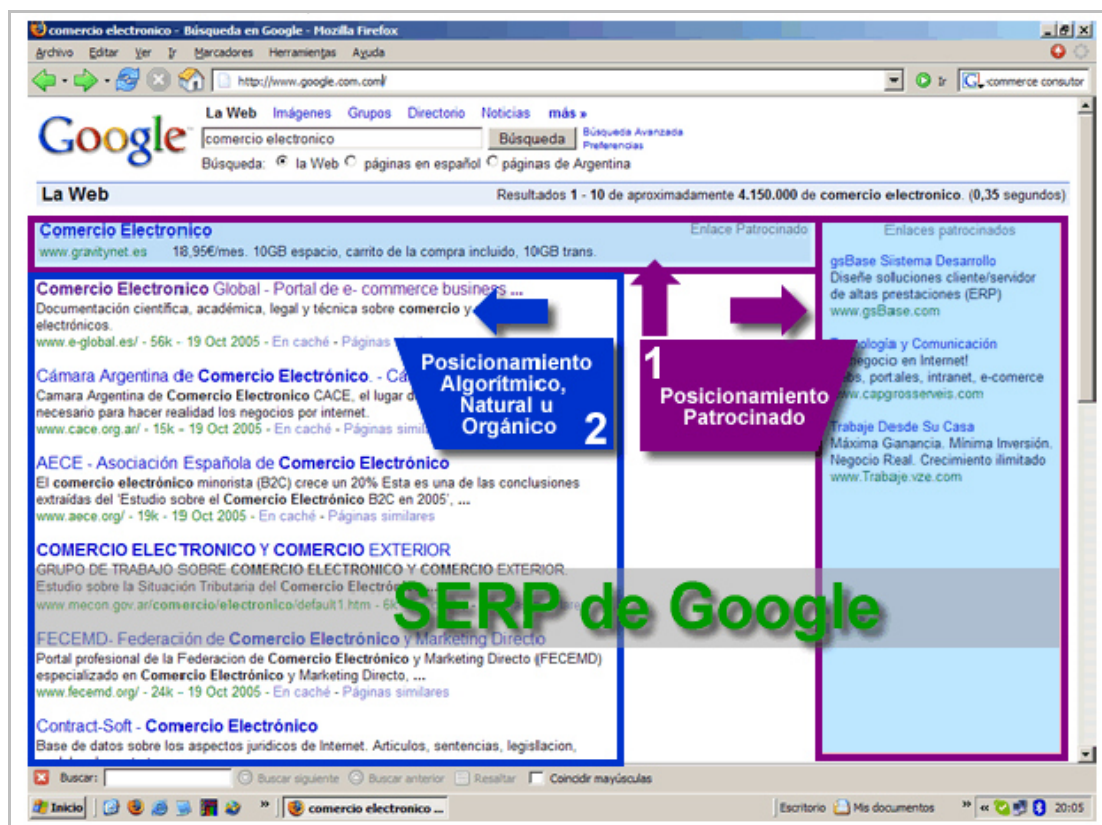


Figura 3.2: Resultados Naturales y Resultados Patrocinados en Google

De ahora en adelante, cuando se mencione posicionamiento en motores de búsqueda haremos referencia al tipo de posicionamiento natural, el cual está basado en los algoritmos de los motores de búsqueda.

3.3 Posicionamiento en Motores de Búsqueda

Debido a que la presente investigación se basa en los resultados naturales de una búsqueda, se centrará en el posicionamiento en Motores De Búsqueda, el cual es el posicionamiento de resultados naturales.

El posicionamiento en Motores de Búsqueda o posicionamiento natural, consiste en un conjunto de procedimientos y técnicas que tienen como finalidad dotar a un portal o a una página Web de la máxima visibilidad en Internet, es decir busca optimizar una página Web para que aparezca en las primeras posiciones de las páginas de resultados de los motores de búsqueda [15].

El posicionamiento en Motores de Búsqueda, también puede ser considerado como arte y técnica. Cada buscador utiliza un algoritmo de indexación diferente que es permanentemente actualizado y mejorado. Estos algoritmos son mantenidos en secreto por las empresas que los crean. Esto hace que el posicionamiento sea un auténtico ejercicio de ingeniería inversa [16].

Una buena estrategia de posicionamiento, abarca todo un proceso de publicidad Web en Internet, utilización de técnicas de Marketing de palabras y frases claves, así como análisis y estudio de la competencia.

3.4 Las fases de un Proceso de Posicionamiento en Motores de Búsqueda

El posicionamiento en motores de búsqueda es un proceso dinámico que tiene fases determinadas en su ciclo de vida, pero a diferencia de otros, no tiene un final ya que cualquier cambio en el portal Web o en los algoritmos de los buscadores, hacen necesario una monitorización y actualización del proceso de posicionamiento de una Web.

A continuación se lista las fases de un proceso de posicionamiento en Motores de Búsqueda:

1. Análisis del portal Web
2. Diseño de las estrategias
3. Implementación de las medidas adoptadas
4. Seguimiento de los resultados, esta fase se aplica pasado un tiempo (al menos un mes) de la aplicación de la fase anterior.

Es necesario recalcar que deberá aplicarse el ciclo de vida de forma individual para tantas secciones de interés como tenga el portal Web. La razón es la siguiente: dada la naturaleza de la Web, no hay razón por la cual se pueda decir que los usuarios accederán todos a través de la página principal del Portal Web. Además, si un portal Web es rico en contenidos, se supone que estaremos interesados en promover las distintas secciones de modo independiente y no solamente la página principal [5].

3.4.1 Fase de análisis

El objetivo de la fase de análisis, es presentar propuestas de optimización específicas en los siguientes aspectos:

- Elección de las palabras claves.
- Identificación de los posibles sitios que pueden enlazar al portal.

Las palabras claves son términos de consulta que utilizan los usuarios de los motores de búsqueda. Además, se denominan palabras claves a los términos presentes en el contenido de un portal que, eventualmente, pueden coincidir con las palabras usadas por los usuarios cuando realizan sus búsquedas.

Una palabra clave puede ser una palabra simple; por ejemplo, “Arte” o una expresión compuesta; por ejemplo, “Historia del Arte”. Puede ser un nombre común o un nombre propio. Puede expresar conceptos o cosas tangibles. También pueden corresponder a ramas del conocimiento (“arquitectura”) y actividades humanas (“reparación de ordenadores”).

Naturalmente, una palabra clave también es el término con el que se representa una necesidad de información; por lo tanto, al seleccionar bien las palabras clave, se estará haciendo de cara a las preguntas que harán los futuros usuarios o el futuro público del sitio [17].

La selección de palabras claves están en función de los tres elementos principales que definen a un portal Web: el contenido, los objetivos y el público, haciendo énfasis en este último elemento, es decir el público al que está destinada el portal.

Para una mejor selección, se recomienda comenzar por buscar tres o cuatro palabras clave principales relacionadas con el portal. Por ejemplo, en el portal Web de un museo de cinematografía, podrían ser las palabras “museos” y “cine”; y después derivar de ellas cinco o seis palabras clave adicionales más específicas. Por último, deben añadirse nombres propios; por ejemplo, el nombre de la institución.

En esta tarea es importante involucrar al departamento comercial en caso de que se trate del sitio de una empresa. En concreto, para ayudar a la determinación de las palabras

clave existen diversos procedimientos, los cuales veremos con más detalle en el siguiente capítulo.

3.4.2 Fase de diseño

El objetivo de la fase de diseño es presentar propuestas de optimización específicas en los siguientes aspectos:

- Optimización de la Web: Estructura, Textos, Enlaces internos, Mapas de navegación, Elección de los títulos, Metatags, Tags semánticos, Texto alternativo de imágenes, scripts y objetos.
- Recomendaciones de cambios en estructura de la Web: establecimiento de una estrategia de URLs, redacción de textos y metas.
- Generación de archivos para Google Site Maps (creación de archivos xml como por ejemplo sitemap.xml).
- Realización de los cambios en el código de la página.
- Optimización de contenidos del portal Web: planificación de la creación regular de contenidos de calidad para publicar en el portal.

3.4.3 Fase de implantación

El objetivo de la fase de implementación es presentar propuestas de optimización específicas en proveer de enlaces de entrada: captación de enlaces corporativos y de sitios de temática similar, por el cual puedan localizar un portal Web.

3.4.4 Fase de seguimiento

El objetivo de la fase de seguimiento es presentar propuestas de optimización específicas en los siguientes aspectos:

- Calendario: monitorizar resultados, los motores de búsqueda pueden tardar semanas en indexar el sitio desde que recibieron el alta.
- Constatar indexación: conocer si el sitio figura en su índice y cuantas páginas han sido indexadas.
- Búsquedas por las diversas palabras clave para las cuales se ha optimizado el portal y anotar en qué posición queda en la página de resultados de cada motor.

El objetivo debe ser aparecer en las primeras páginas de resultados y, a ser posible, en la primera.

3.5 Ventajas de un Buen Posicionamiento en Motores de Búsqueda

- **Tráfico cualificado**

La ventaja de un portal Web bien posicionado es que está dirigido a los usuarios interesados en los productos y servicios que ofrece dicho portal. El usuario elige libre y voluntariamente el buscador, las palabras clave y los resultados de la S ERP.

- **Aumento de la Imagen Corporativa**

La ventaja de un portal Web bien posicionado es que aumenta la visibilidad de la página para diferentes palabras claves. Percepción de su empresa como Referente del Sector (Importancia en el Campo en que se mueve su empresa).

- **Bajo costo**

El tráfico obtenido a través de una campaña constante de posicionamiento en buscadores tiene el coste por visitante real más bajo que cualquier otro método de promoción por Internet.

- **Durabilidad**

Una campaña de posicionamiento en Motores de Búsqueda que fueron bien diseñados sigue teniendo efectos muchos meses, e incluso años, después de implementarla.

Capítulo 4: Técnicas Existentes para el Posicionamiento en los Motores de Búsqueda

En el capítulo anterior se vio acerca del Posicionamiento en Buscadores como una forma de promoción Web. Ahora, en este capítulo, daremos a conocer las principales técnicas que existen para lograr un buen posicionamiento en los principales Motores de Búsqueda, como son Google, Yahoo y MSN.

4.1 Introducción

Como se mencionó anteriormente, desarrollar y publicar un portal Web es un paso importante pero no suficiente, también es necesaria una campaña de promoción Web para poder generar grandes cantidades de visitas y oportunidades de negocio.

Es así que en [5], se describe que hay muchas formas de promoción Web que pueden atraer a clientes potenciales y generar más oportunidades de negocio. Entre las principales están:

- Posicionamiento publicitario.
- Posicionamiento en motores de búsqueda.
- Publicidad a través de banners.
- Los envíos de correo masivos.
- El marketing de bitácoras (blogs).
- Directorios especializados.

Podemos destacar a las 2 primeras formas, no sólo porque tiene que ver con el tema de investigación (posicionamiento en buscadores), sino también porque, los buscadores son la herramienta por excelencia para localizar información en la Internet, y se han convertido en el medio más efectivo para promocionar productos y/o servicios.

En ese sentido, en las siguientes secciones detallaremos acerca de las técnicas existentes para conseguir un mejor posicionamiento en estos motores de búsqueda.

4.2 Pago por Clic (PPC)

En ingles llamado *Page per Clic*. Es una técnica rápida para posicionar un Portal Web consiguiendo enlaces patrocinados con lo cual, se genera tráfico a través de los motores de búsqueda.

4.2.1 Enlaces Patrocinados

Complementando lo que se dijo en la sección 3.2, acerca de los enlaces patrocinados, son anuncios situados en los SERPs, pero en una zona claramente diferenciadas de los resultados orgánicos o resultados naturales. Este tipo de publicidad es de pago y se factura por cada clic que los usuarios del buscador hagan en el enlace patrocinado que conduce al portal Web patrocinado.

Localizar un portal Web en la zona de “Enlaces patrocinados”, es el objetivo del Posicionamiento Publicitario, el cual es una de las opciones para promocionar mejor un portal Web.

Si la empresa desea optar por el Posicionamiento Publicitario, puede hacerlo mediante convenios de Pago por Clic (PPC) con los principales buscadores mediante sus servicios de publicidad.

A continuación se muestra los servicios publicitarios de mayor uso:

- Servicio Adwords de Google (tiene el 90% del mercado).
- Servicio Sponsored Search de Yahoo.
- Servicio Microsoft Digital Advertising Solutions de Live.

Según [5], Google, Yahoo y Live son los que se reparten el mayor tráfico de búsqueda, con más del 90% de las búsquedas realizadas. Según los últimos datos de enero del 2008, Google tendría el 60% y el 65%, Yahoo el 20% y Live un 10%.

Estos porcentajes se reflejan también en la cuota de mercado en publicidad en buscadores que también está a favor de Google pero con mayor diferencia. Debido a que Google Adwords fue el primer servicio de publicidad en buscadores con altas prestaciones, los datos que se manejan actualmente son que el 90% del mercado está en manos de Google.

4.2.2 Campaña PPC

Una campaña de PPC viene a ser una manera rápida de generar tráfico a través de los motores de búsqueda. Muchas compañías, particularmente compañías grandes con grandes presupuestos de marketing, aplican directamente PPC y pasan por alto la técnica SEO (tratado en la sección siguiente).

La utilización de PPC incluye pagar cada vez que alguien hace clic en uno de los anuncios. Existen varias clases de estos anuncios, donde se paga por cada clic que hacen en ella, pero hace solo dos años estos enlaces fueron incluidos en los resultados de los motores de búsqueda [18].

En la figura 4.1, se ve como se muestra los resultados patrocinados de una búsqueda en Google para la palabra “hotel Perú”, y en la figura 4.2, se ve como muestra los resultados patrocinados para la misma palabra de búsqueda en Yahoo. Como se puede ver, muchos de los resultados de la búsqueda en estas páginas son realmente anuncios y se colocan basados principalmente en una tarifa. Todo esto muestra el apuro y la necesidad de las empresas por optimizar su página Web y conseguir enlaces de otros sitios a su portal.

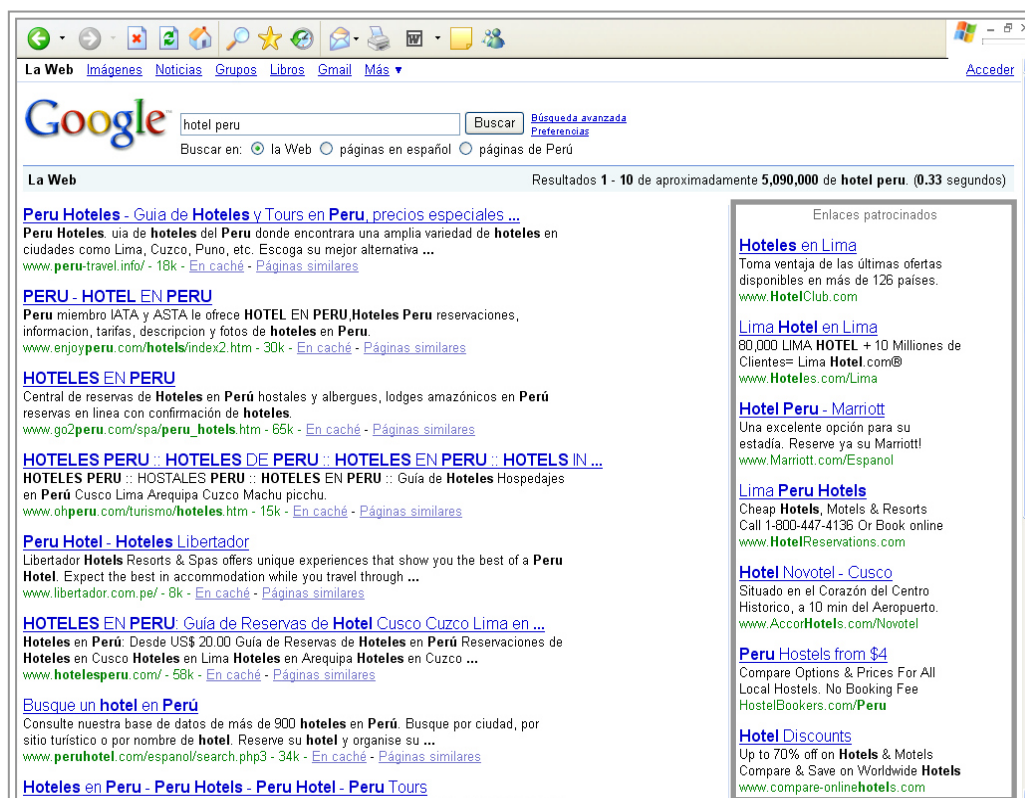


Figura 4.1: Enlaces Patrocinados para una Palabra de Búsqueda en Google.

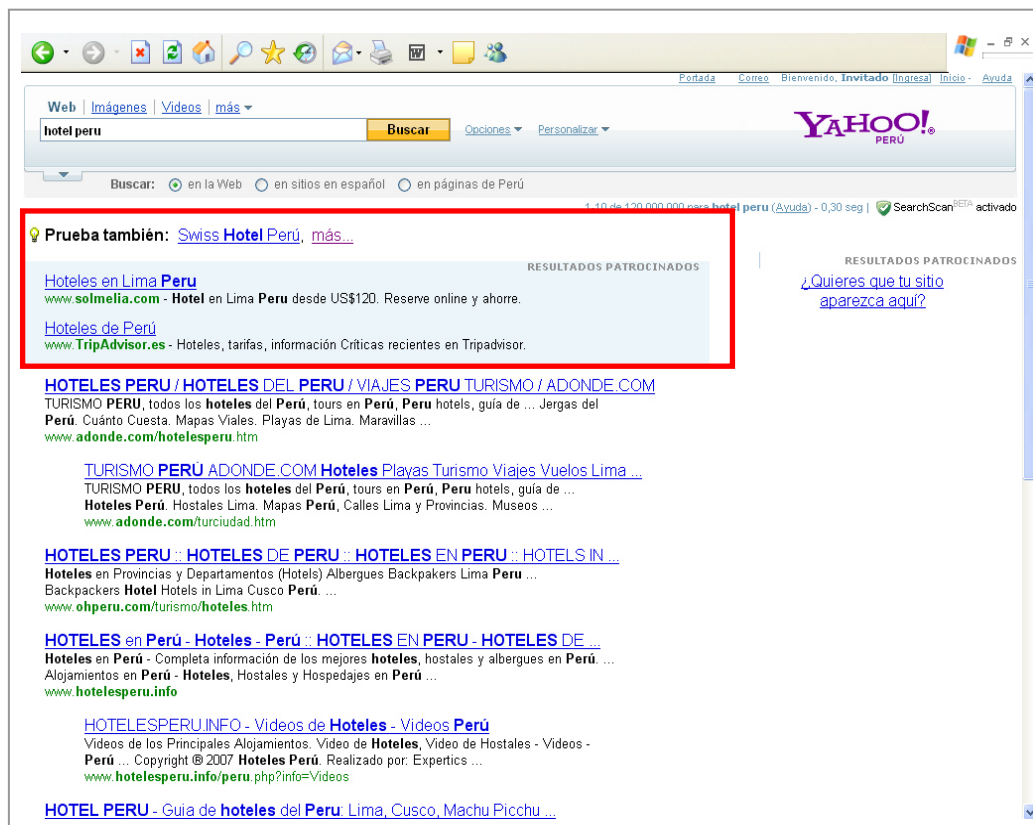


Figura 4.2: Enlaces Patrocinados para una Palabra de Búsqueda en Yahoo

4.2.3 Funcionamiento de una Campaña PPC

Peter Kent en [18], menciona como trabaja una Campaña PPC en los principales Motores de Búsqueda, los cual se resume a continuación:

1. Registrarse en el sistema del PPC, se debe proporcionar un número de tarjeta de crédito.
2. Proceder a crear uno o más anuncios, se debe proporcionar un title, body text y un enlace a la página Web que se quiera que se dirijan los usuarios visitantes.
3. Asociar palabras claves a cada anuncio.
4. Hacer una oferta en cada palabra clave.

En la figura 4.3, se diagrama como se puede representar el flujo de una subasta de anuncios (oferta de una palabra clave).

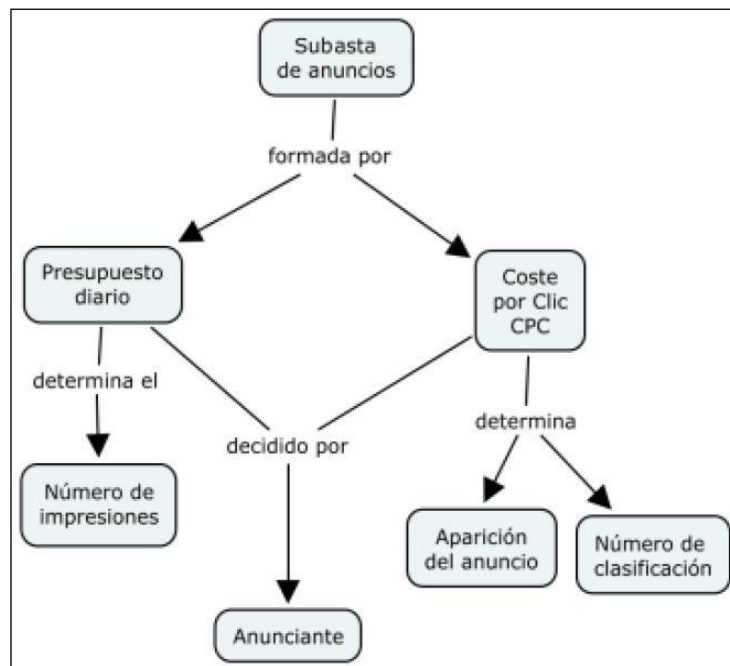


Figura 4.3: Representación del Flujo de una Subasta de Anuncios.

Generalmente, cuando se hace una oferta en una palabra clave, se hace una oferta en una posición. En la figura 4.4, se ve una lista de palabras claves en el sistema del PPC. La primera entrada demuestra cuánto se está dispuesto a ofrecer en la palabra clave “academic software”. Es decir si alguien busca con “academic software”, se pagará \$1.51 si esa persona hace clic en su anuncio. Y, en este caso, debido a que otras empresas ofrecen pagar menos por la misma palabra clave, su anuncio se mostrará primero. Se puede ver las otras ofertas en la columna “Top 5 Max Bids” [18].

Search Term	Broad/Phrase	Position	Your Cost	Your Max Bid	Top 5 Max Bids					Click Index™	Impressions	Clicks	Click Rate	Average Cost	Total Cost
				Update Bids							Stats for My Terms: Month to Date				
academic software		1	1.51	1.51	1.51	1.50	1.45	1.45	1.45	More Bids	783	88	11.2	1.49	131.47
activity childhood early learning		1	0.10	0.10	0.10	-	-	-	-	More Bids	11	0	0.0	0.00	0.00
best educational software		1	0.11	0.11	0.11	0.10	0.10	0.10	-	More Bids	12	2	16.7	0.11	0.22
child educational game		1	1.59	1.59	1.59	1.58	0.30	0.27	-	More Bids	570	13	2.3	1.59	20.67
child educational software		2	0.44	0.70	0.70	0.70	0.43	0.42	0.41	More Bids	127	3	2.4	0.44	1.32
childhood early learning music		1	0.10	0.10	0.10	-	-	-	-	More Bids	0	0	0.0	0.00	0.00
childhood early learning toy		1	0.10	0.10	0.10	-	-	-	-	More Bids	1	0	0.0	0.00	0.00
early childhood learning		1	0.13	0.13	0.13	0.12	0.10	-	-	More Bids	24	5	20.8	0.13	0.65
early childhood learning supply		1	0.10	0.10	0.10	-	-	-	-	More Bids	0	0	0.0	0.00	0.00
educational adventure software		1	0.10	0.10	0.10	-	-	-	-	More Bids	0	0	0.0	0.00	0.00
educational computer game		1	0.11	0.11	0.11	0.10	0.05	0.05	-	More Bids	87	7	8.0	0.11	0.76
educational computer		1	1.02	1.02	1.02	1.01	1.00	0.27	0.26	More Bids	76	8	10.5	1.02	8.16

Figura 4.4: Ofertas de Palabras Claves en una Campaña PPC

En la mayoría de los sistemas del PPC, la mayor oferta consigue el lugar más alto. Pero en el sistema de Google es un poco más complicado, no se puede estar exactamente seguro en que posición aparecerá el anuncio. Cuando Google coloca un anuncio en el resultado patrocinado de una búsqueda, elige la posición basada no sólo en el precio de oferta sino también en el ratio de clickeo (*click through ratio*, *CTR*), el cual representa la frecuencia con la que los visitantes hicieron clic en el anuncio en el pasado.

Se puede ofrecer más que cualquier otro, pero si se tiene un bajo CTR, es probable que el anuncio respectivo sea colocado por debajo del que ofreció menos. Por eso Google dice: “*Los anuncios más relevantes se mostrarán en el lugar más alto...*”, como se menciona en [18]. Es así que un anuncio puede colocarse sobre alguien que paga más si es que es altamente relevante para una palabra clave específica.

Ya que la posición de un anuncio por PPC depende en parte de la oferta que se hace, aparece una pregunta importante ¿Cuánto se debe ofrecer? A continuación se muestra algunas cosas a considerar:

- La empresa debe entender cuánto es realmente el valor de un clic en su anuncio.
- Mientras sea más alta la posición del anuncio, hay más probabilidad de conseguir un mayor tráfico.
- Generalmente, si no se está en las primeras tres posiciones, hay una buena probabilidad de que el anuncio no será visto a menudo y conseguirá bajar en el CTR.
- A veces se puede querer tomar la posición tres, ya que se quiere estar entre los tres primeros, debido a que la primera posición es demasiado costosa.

Overture, Google AdWords y algunos otros, ajustan automáticamente su precio de oferta de modo que se pague solamente \$0.01 sobre el licitador más bajo, es decir, por ejemplo, si han hecho una oferta \$3.50 para una frase de la palabra clave, mientras que el licitador más alto siguiente ha hecho una oferta solamente \$2.99. Le cargarán \$3.00 para un tecleo. Si el licitador siguiente levanta su oferta a \$3.01, le cargarán \$3.02, y así sucesivamente. Sin embargo, observe que algunos sistemas más pequeños no hacen esto, así que se tiene que mirar las ofertas cuidadosamente o se termina pagando más de lo que se necesita.

También, Peter Kent, expone que la mayoría de los sistemas de PPC tienen dos tipos de anuncios: anuncios del Motor de Búsqueda y anuncios contextuales. El primer tipo ya se vio líneas arriba con los “enlaces patrocinados”.

En cambio, los anuncios contextuales son anuncios colocados en diferentes portales Web. Por ejemplo, en Amazon.com aparecerán probablemente anuncios de acuerdo a alguna búsqueda, en la figura 4.5, se visualiza un anuncio contextual con la palabra “joyería” (http://www.amazon.com/s/ref=nb_ss_gw?url=search-alias%3Daps&field-keywords=joyeria).

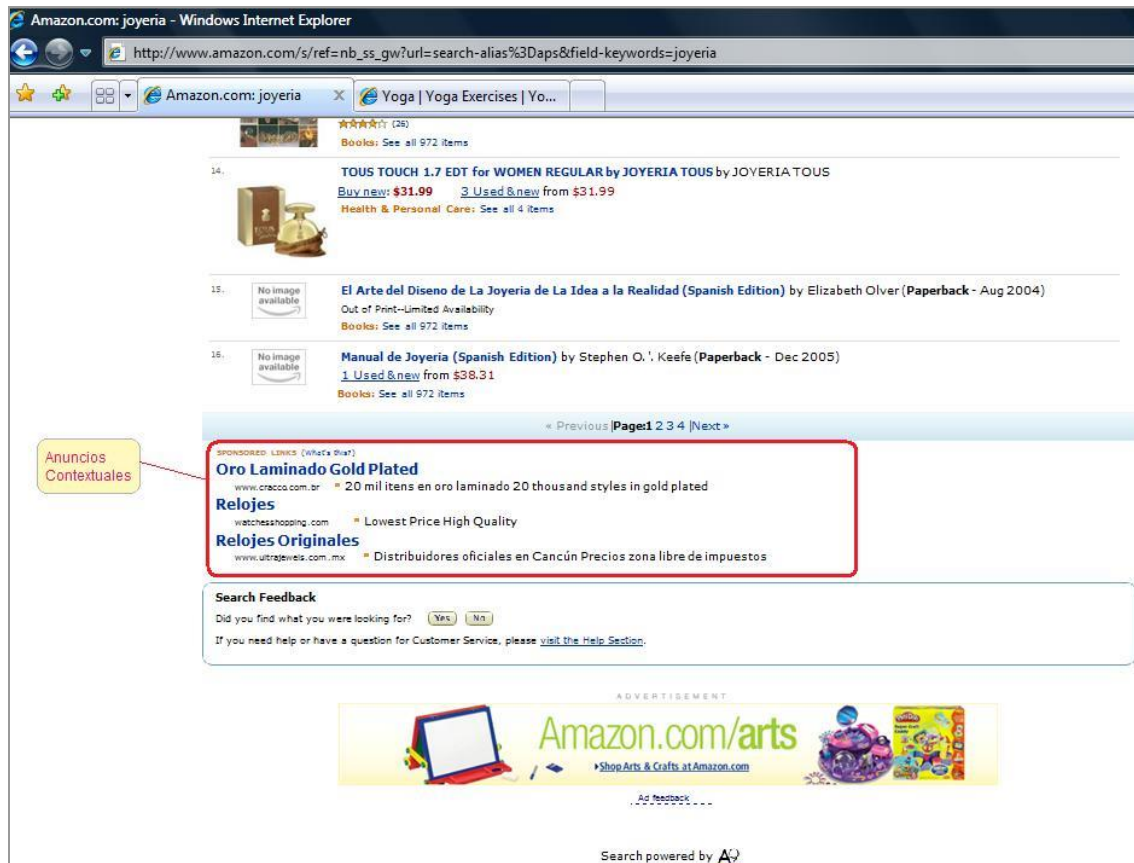


Figura 4.5: Anuncios Contextuales en Amazon.com

4.2.4 Ventajas de una Campaña PPC

Las ventajas de llevar a cabo una campaña PPC para mejorar el posicionamiento de un portal Web en los Motores de Búsqueda, vienen a ser que:

- Representa una manera mucho más rápida de comenzar a generar tráfico al portal Web, es cuestión de horas o días, comparados a las semanas o a los meses en una búsqueda natural.
- Permiten una flexibilidad absoluta y casi instantánea en su gestión, pudiendo manejar rápida y eficazmente y hacer cambios en sus campañas en tiempo real.

- Tienen una mayor probabilidad de conseguir tráfico, no importa si no se esfuerza en optimizar las páginas; con el PPC se conseguirá lo que se paga. Además, si se está dispuesto a hacer una oferta bastante grande, se conseguirá mayor tráfico.
- Es más estable. Un portal Web puede estar bien hoy en los resultados naturales pero eso quizás cambié en un día. Con el PPC, mientras se esté dispuesto a pagar, el portal seguirá siendo mostrado.
- Una campaña PPC puede ayudar a lograr metas grandes, y una buena campaña PPC es mucho más barato que cualquier forma de comercialización fuera de la Internet.

4.2.5 Desventajas de una Campaña PPC

Por otro lado, las desventajas de llevar a cabo una campaña de PPC son las siguientes:

- Se tiene que pagar por cada clic, que puede significar mucho dinero.
- En el caso de las pequeñas y medianas empresas, el problema del costo se acrecienta. Será más difícil para ellos, acceder a este tipo de campaña.
- Una campaña PPC no funciona siempre, no es posible que cada empresa compre clics en un precio que sea bajo y bastante provechoso.
- Mucha gente no hace clic en los enlaces patrocinados de la búsqueda, pues no confía en ella [18].
- Hay rubros (mercados de negocio) muy competitivos, donde los clics para las palabras clave más utilizadas por los usuario, son muy caros y se hace difícil publicar en estos rubros.

En conclusión, la técnica PPC podría formar parte de una campaña de marketing total. Sin embargo, debe combinarse con las otras técnicas existentes. Muchas compañías pagan sumas enormes en campañas PPC debido a que ignoran totalmente la búsqueda natural.

4.3 Técnica SEO

Es otra forma de promocionar un portal Web y lograr un buen posicionamiento en los principales Motores de Búsqueda, a fin de conseguir más visitas. Es aplicar la Técnica de Optimización para Motores de Búsqueda (*Search Engine Optimization* – SEO), el cual no es otra cosa que la mejora del posicionamiento orgánico o natural en los SERPs. Los resultados se muestran sin previo pago, a diferencia de los resultados publicitarios o enlaces patrocinados.

Además en [19], el posicionamiento en motores de búsqueda de los resultados orgánicos, depende de más de 100 parámetros que recogen los algoritmos que utilizan los robots de los buscadores, para encontrar este recurso entre los millones que tienen indexados. Los parámetros que utilizan los diferentes motores de búsqueda no son conocidos, ya que forman parte de su ventaja competitiva.

En ese sentido, según Pablo Lara y José A. Martínez, la Técnica SEO viene a ser un conjunto de procedimientos cuya finalidad es mejorar el posicionamiento de un portal o página Web en los SERPs.

4.3.1 Optimización para Motores de Búsqueda(Search Engine Optimization– SEO)

Según Jennifer Grappone y Gradiva Couzin en [20], el término SEO describe la optimización para motores de búsqueda, siendo un conjunto diverso de actividades que se puede realizar para aumentar el número conveniente de visitantes que vienen a un sitio Web a través de los motores de búsqueda.

También, Kristopher B. Jones en [21], sostiene que SEO es el proceso de configurar un sitio Web de modo que pueda ser bien posicionado para determinadas palabras claves dentro de los resultados orgánicos de una búsqueda en los motores de búsqueda más importantes como Google, Yahoo y MSN. A diferencia de los procesos de Marketing en Buscadores, donde se debe pagar por cada clic que enlace al sitio Web, el tráfico enviado a su sitio desde los resultados orgánicos de un Motor de Búsqueda son gratis.

Por consiguiente, podemos definir a SEO como un proceso que mejora el volumen y la calidad de tráfico vía “natural”, por el cual se pretenden conseguir posiciones elevadas

en los Motores de Búsqueda, para un conjunto de palabras o frases clave determinadas. Es decir, corresponde a las acciones a tomar para mejorar el resultado que se obtiene en los buscadores sin pagar directamente al buscador. Esto no quiere decir que el trabajo SEO vaya a salir gratis, pero sí es más económico que PPC, en relación con el número de visitas que proporciona.

Adicionalmente en [20], se menciona que SEO abarca actividades de seguimiento, investigación, revisión de la competitividad.

4.3.2 Criterios Básicos (Recomendaciones) para conseguir un Buen Posicionamiento

Los Motores de Búsqueda mantienen en secreto el detalle último de sus procedimientos, ya que se trata de una información que puede dar ventajas competitivas frente a otros y, por tanto, la consideran un secreto industrial.

Por este motivo, todo aquello que los estudiosos y profesionales afirman sobre el tema, en realidad es, o bien simple especulación, o bien resultado de inferencias indirectas. Es decir, a partir de la observación y del análisis de los resultados, existe un cierto consenso entre los analistas sobre qué clase de criterios usan los motores de búsqueda para ordenar los resultados [19].

Estos consensos sobre los criterios que usan los motores de búsqueda son tomados para definir los procesos y/o prácticas que se debe seguir para conseguir un buen posicionamiento.

A continuación se especifican algunos criterios que se debe seguir en una exitosa aplicación de técnicas SEO, las cuales están basadas en los criterios de la mayor parte de los analistas siguen los tres principales motores de búsqueda (Google, Yahoo y MSN).

Ahora bien, aunque toda la evidencia apunta hacia el hecho de que los criterios señalados a continuación son los más importantes, se ignora como combinan, en cada momento, la importancia de cada uno de ellos. Además, tales criterios pueden variar a lo largo del tiempo.

Miguel López Gómez, en [22], clasifica a estos criterios que influyen en la posición de los Motores de Búsqueda, en 2 grandes grupos (factores):

- Factores Internos.
- Factores Externos.

4.3.2.1 Factores Internos

Que depende exclusivamente del propietario del Portal Web y su equipo de desarrollo, del contenido de las páginas Web, así como de los metadatos utilizados para la descripción de las mismas y la manera en que se estructura dicho portal Web.

A continuación, se abordará los diferentes criterios internos para conseguir un buen posicionamiento en los principales Motores de Búsqueda (Google, Yahoo y MSN).

a. Optimización de las Palabras Clave

iii. Definición

Empezaremos definiendo que son las palabras claves, se llama palabra clave al término de consulta que usan los usuarios en los motores de búsqueda (véase la figura 4.6), así como a los términos presentes en el contenido de un portal Web que, eventualmente, pueden coincidir con las palabras usadas por los usuarios cuando realizan sus búsquedas (véase la figura 4.7).



Figura 4.6: Palabra Clave “arte” – Término de Consulta en Google



Figura 4.7: Palabra Clave “arte” – Termino Presente en el Contenido del Portal Arte y Cultura del Perú (<http://www.arteperu.com>).

Una palabra clave puede ser una palabra simple; por ejemplo, “Arte” o una expresión compuesta; por ejemplo, “Historia del Arte”. Puede ser un nombre común o un nombre propio. Puede expresar conceptos o cosas tangibles [18].

Para Kristopher B. Jones, hay 2 clases de términos de palabra clave (*Keywords*):

- Términos Principales (head terms), refieren a las palabras clave cortas, generalmente una a dos palabras en longitud que tiendan a tener volumen muy alto de la búsqueda.
- Términos de Cola (tail terms), los términos de cola son los términos más largos de la palabra clave, típicamente tres o más palabras que tienden a tener un volumen mucho más bajo de la búsqueda que los términos principales.

iv. Importancia

En el capítulo 4, “*Picking Powerful Keywords*”, de [18], Peter Kent menciona que la selección óptima de las palabras clave es la base de toda estrategia SEO, muchas veces lograr que los Motores de Búsqueda reconozcan e indexen un portal Web no tiene nada que ver con las limitaciones tecnológicas de estos motores, sino con la elección de las palabras clave.

A continuación se presenta un ejemplo que muestra, como los motores de búsqueda utilizan las palabras clave:

Un usuario va a un motor de búsqueda y trata de encontrar algo, escribe “Desarrollo Sistemas”, para finalmente hacer clic en el botón Buscar.

A continuación, el motor de búsqueda busca en sus índices la palabra o palabras ingresadas, luego mostrará:

- Páginas que contengan la frase exacta de “Desarrollo Sistemas”.
- Páginas que no cuentan con la frase de “Desarrollo Sistemas”, pero tienen las palabras “Desarrollo” y “Sistemas”.
- Páginas que tienen las palabras “Desarrollo” y “Sistemas”, aunque no necesariamente juntos.
- Páginas con las palabras que derivan de “Desarrollo” y “Sistemas” en algún lugar de la página.
- Las páginas que contienen vínculos que contiene el texto del vínculo, la frase de “Desarrollo Sistemas”.
- Páginas con vínculos que apuntan a ellas con el enlace de texto que contienen las palabras “Desarrollo” y “Sistemas”, aunque no juntos.

El proceso es en realidad mucho más complicado que esto. El motor de búsqueda no necesariamente indica las páginas Web en el orden mencionado. Más bien, al examinar el orden para clasificar las páginas, el Motor de Búsqueda tiene en cuenta otras características de la palabra o palabras clave:

- ¿Las palabras clave se encuentran en negrita?
- ¿Las palabras clave se encuentran en cursiva?

- ¿Las palabras clave se encuentran en las listas con viñetas?
- ¿El texto de las palabras clave son más grande que la mayoría de los otros textos en la página?
- ¿Las palabras clave se encuentran en el heading text (<H> etiquetas)?

Y otros criterios más. A pesar de todas las diversas complicaciones, sin embargo, un hecho que es de suma importancia: Si un buscador no puede enlazar a un sitio Web con las palabras que busca alguien, no tiene ninguna razón para volver a ese sitio Web como parte de los resultados de búsqueda.

Recoger las palabras clave adecuadas es fundamental. Como Woody Allen dijo una vez: *"Eighty percent of success is showing up"*, El ochenta por ciento de éxito es ser visto. "Si no se juega no hay oportunidad de ganar". Y si usted no elige las palabras clave adecuadas, no será visto y ni siquiera podrá jugar el juego.

v. Problema

El principal problema con las palabras claves es que muchas empresas las descubren luego de crear su portal Web, y si ya tenían sus palabras claves en el portal, descubren que las palabras claves usadas por sus potenciales clientes y visitantes no son las mismas que se está utilizando en el portal.

Este se debe a 2 posibles razones:

- Nuevamente encontramos el problema de que la mayoría de los sitios Web se construyen sin tener en cuenta los motores de búsqueda. O tienen poco conocimiento acerca de cómo trabajan éstos.
- Y si el sitio Web fue diseñado pensando también en los motores de búsqueda, a menudo se usa palabras clave incorrectas.
- Un viejo concepto afirma que se debe pensar como sus clientes. Las empresas a menudo cometen errores con sus palabras, ya que recogen las palabras clave de acuerdo a como ellos (en lugar de sus clientes) piensan en sus productos o servicios.

vi. Como Optimizar

La solución es poder analizar que palabras pueden ser claves para el portal Web, es decir, aquellas que describan mejor el portal. También se debe descubrir cuales son las palabras potencial que los visitantes del sitio utiliza para buscar los productos y servicios del portal Web.

Las palabras clave serán elegidas en función de los tres elementos principales que definen a un portal Web: el contenido, los objetivos y el público (usuario), haciendo énfasis en este último elemento, es decir el público al que está destinado el portal Web.

- SEO recomienda comenzar por buscar tres o cuatro palabras clave principales relacionadas con el portal y después derivar de ellas cinco o seis palabras clave adicionales más específicas. Además SEO recomienda añadir nombres propios; por ejemplo, el nombre de la institución.
- SEO recomienda buscar aquellas palabras con las que se quiere que determinado portal Web obtenga altas posiciones en las páginas de resultados de los buscadores.
- Debido a que una palabra clave también es el término con el que se representa una necesidad de información; por lo tanto, se debe seleccionar bien las palabras clave, para así hacer frente a las preguntas que harán los futuros usuarios o el futuro público del portal Web.
- SEO recomienda hacer algunas investigaciones para averiguar como los consumidores realmente llaman a sus productos. Hacer un poco el análisis de palabras clave, para saber y comprobar lo que la gente realmente busca en el portal Web.

- Recordar que las palabras claves específicas, que incluyen esas palabras claves que describan su producto o servicio específico y sean más de tres palabras claves en longitud, son menos competitivas y por lo tanto deben componer el bulto de sus esfuerzos de la generación de la palabra clave.
- SEO recomienda realizar una mezcla de los términos principales y de los términos de cola.

Una de las herramientas para generación de palabras clave más usadas, previa suscripción y paga, es Wordtracker.com, provee una lista de términos relacionada a la palabra clave base, así como el volumen diario estimado de la búsqueda para cada palabra clave individual.

b. Optimización de los Títulos <TITLE>

Uno de los aspectos más importantes a tener en cuenta en una estrategia SEO es conocer los criterios que utiliza el motor de búsqueda en relación con las diferentes partes de un documento HTML.

La estructura típica de un documento HTML muestra diferentes partes, tal como se muestra a continuación [16]:

```
<HTML>
<HEAD>
    <!-- información de cabecera utilizada para contener información
    adicional acerca del documento, y que no se muestra en pantalla -->
</HEAD>
<BODY>
    <!-- todo el HTML a visualizar -->
    ::
    ::
    ::
</BODY>
</HTML>
```

Una de las etiquetas es la <TITLE>, La cual se encuentra dentro del <HEAD>, transmite al robot del motor de búsqueda el título de la página. Dado que la etiqueta <TITLE> de una página Web es lo que los buscadores muestran en la lista de resultados, el objetivo de la optimización es doble. Por un lado debe ser una manera de llamar la atención, para que los usuarios entren al portal Web, y por otro, debe estar orientado de tal forma que los motores de búsqueda otorguen una buena posición a dicho portal.

Entre las prácticas recomendadas por SEO para las etiquetas <TITLE> están:

- **Describir con precisión el contenido de la página.** Elegir un título que indique claramente el tema de la página.

Evitar títulos que no tengan ninguna relación con el contenido de la página, el uso de títulos predeterminados o demasiado genéricos como "Sin título" o "Página nueva 1".

- **Crear etiquetas <TITLE> únicas para cada página.** Lo ideal es que cada una de tus páginas tenga una etiqueta <TITLE> única, que ayude a Google a distinguir esa página del resto de páginas de tu sitio.

Evitar el uso de una sola etiqueta <TITLE> para todas las páginas de tu sitio, o para muchas de ellas.

- **Usar títulos descriptivos breves.** Los títulos pueden ser concisos pero informativos. Si el título es demasiado largo, Google mostrará tan solo una parte del mismo en el resultado de búsqueda.

Evitar títulos muy largos, que no sean útiles para los usuarios, rellenar las etiquetas <TITLE> con palabras clave innecesarias.

c. Las Metaetiquetas o Metadatos

Los lenguajes de marcado (forma de codificar un documento) como html, xhtml, dhtml y otros permiten utilizar una serie de etiquetas, denominadas Meta, a través de las cuales se puede añadir una serie de informaciones sobre una página. Principalmente se suelen utilizar para describir el contenido a través de pequeños resúmenes y palabras-clave. Hay robots que son capaces de identificar las etiquetas META y extraer la información de las mismas para ser usada en la búsqueda o en la presentación de resultados.

Los metadatos están incluidos dentro de la etiqueta <HEAD>, tienen como objetivo ofrecer a los buscadores información acerca del recurso electrónico. Las más importantes, luego del <TITLE> comentado más arriba son:

- La etiqueta META DESCRIPTION, la cual describe el contenido de la página y sirve de descripción en los resultados de algunos buscadores,
- La etiqueta META KEYWORDS, actualmente casi todos los buscadores la ignoran debido a su manipulación para conseguir mayor relevancia.
- La etiqueta META LANGUAGE, en la cual se indica el idioma de la página.
- La etiqueta META ROBOTS, en la cual se indica al buscador si se desea indexar la página y/o se desean seguir los links.

A pesar de que algunos motores de búsqueda no los tienen en cuenta, se recomienda continuar creándolas para cada una de las páginas de la Web, puesto que suelen ser un factor relacionado con la calidad del recurso y se pueden utilizar para otros propósitos relacionados con la gestión de recursos electrónicos.

Si bien es cierto que se recomienda usar las cuatro metaetiquetas, el que requiere más atención es la META DESCRIPTION, la cual proporciona a Google y a otros motores de búsqueda un resumen sobre la página. Mientras que el título de una página contiene unas pocas palabras, el META DESCRIPTION podría contener un par de frases o incluso un párrafo corto.

Entre las prácticas recomendadas por SEO para la META DESCRIPTION están:

- **Resumir de forma precisa el contenido de una página.** Escribir una descripción que informe y a su vez cree interés en los usuarios en caso de que encuentren esa metaetiqueta description como fragmento de un resultado de búsqueda.

Evita: una metaetiqueta description con contenido no relacionado con la página, descripciones genéricas como "Esto es una página Web" o "Página sobre cromos de fútbol", una descripción con sólo palabras clave, copiar y pegar todo el contenido del documento en una metaetiqueta description.

- **Utilizar descripciones únicas para cada una de las páginas.** Tener una metaetiqueta diferente para cada página ayuda tanto a los usuarios como a

Google, especialmente en búsquedas en las que los usuarios pueden obtener varias páginas de tu dominio (por ejemplo, búsquedas con el operador site:). Si el portal Web tiene miles o incluso millones de páginas, la elaboración de metaetiquetas description a mano no será factible. En este caso, se pueden generar automáticamente basándose en el contenido de cada página.

Evita: utilizar una única metaetiqueta description en todas las páginas de tu sitio o en un gran grupo de páginas de tu sitio.

d. Elementos de Descripción Contextual

Además de los metadatos de la sección head (principalmente title , description y keywords), otras etiquetas también proporcionan información descriptiva de utilidad para los buscadores y donde se deben colocar las palabras clave secundarias:

- Los encabezados, que se codifican mediante <Hn> (donde n es un número del 1 al 6) se utilizan para estructurar jerárquicamente los contenidos principales de una página Web. Por tanto, aquellas palabras clave destacadas deberían situarse en los niveles de encabezado más altos, esto es, H1 y H2.
- El texto de los enlaces, que es la parte activa codificada mediante texto y donde se establecen enlaces internos o externos a los contenidos del sitio Web, contenidos cuyos textos se consideran importantes como palabra clave para la indexación por parte de los motores de búsqueda.
- Los “alt” de las imágenes, según las pautas de accesibilidad a los contenidos Web, se prevé que todas las imágenes deben contener un alt (alternative text o texto alternativo), que debe describir el contenido fundamental de la imagen. Si la imagen no transmite información el atributo alt debe ir vacío.
- Los “title” de los enlaces y las imágenes, en principio la utilización del atributo title para enlaces e imágenes, aunque es valorado por algunos motores de búsqueda puede entrar en contradicción con los criterios fundamentales de la accesibilidad Web. En accesibilidad Web sólo se debe incluir el atributo “title” en un enlace, cuando no es posible activar alguna palabra o palabras significativas. Además, aunque el atributo title se acepta para las imágenes, las Pautas de accesibilidad a los contenidos Web recomiendan la utilización de “alt” y no mencionan, de forma positiva o negativa la utilización del atributo “title”.

- La etiqueta “strong”, que se codifica mediante `texto` y denota que el texto destacado como “strong” tiene cierta importancia. La etiqueta “strong” se utiliza como equivalente a `` (bold), que es una etiqueta desaconsejada en HTML. La etiqueta “strong” tiene mucho peso semántico y se muestra en los navegadores como negrita.
- La etiqueta “em”, que se codifica mediante `texto`, se utiliza para dar énfasis a una palabra o frase, marcando de forma distintiva los puntos más importantes de un texto. La etiqueta “em” tiene menos peso semántico que “strong” y se muestra en los navegadores como cursiva.

e. Palabras Clave Secundarias

Las keywords secundarias se deben distribuir correctamente en el texto de una página. Se recomienda una densidad (porcentaje de palabras clave sobre el total de palabras) de entre el 5% y el 8%. Según el principio del keyword proximity se recomienda situarlas lo más cerca posible del principio de la página. Para darles más importancia existen varias etiquetas: `<Hn>`, `<i>`, ``, `` y ``. La keyword principal se resalta con un `<H1>` y debe colocarse cerca del principio de la página.

4.3.2.2 Factores Externos

Que requieren de la participación de otros portales Web o servicios externos para mejorar su posicionamiento.

a. El PageRank

Es un valor entre 1 y 10 que depende de la cantidad y calidad de los sitios Web que tengan links hacia el sitio Web de referencia, así como de sus links internos. El PR transmitido por los Sitios Web depende a su vez del PR propio y del número de links salientes que tenga esa página [2]. La fórmula del PR es la siguiente: $PR(A) = (1-d) + d \left(\frac{PR(T1)}{C(T1)} + \dots + \frac{PR(Tn)}{C(Tn)} \right)$ PR(A) es el PageRank de la página de referencia. d es un factor de debilitación. (1-d) asegura que cualquier página, aunque no reciba ningún enlace, tendrá un PR mínimo de 0'15. $PR(Ti)/C(Ti)$ es el PageRank (PR) de la página i-ésima que enlaza al sitio Web de referencia, (Ti), dividido por todos los enlaces (C) que también salen de esa página Ti, es decir, el PR que transmite. $i = 1, \dots, n$ ya que se suponen n páginas que enlacen a la de referencia.

b. El Texto de los Links

El texto de los enlaces que apuntan a una página es considerado por algunos como el recurso número uno de la optimización. Las palabras clave se deben colocar en el texto que actúa como anchor de la etiqueta de un link. Agentes inteligentes en la búsqueda y recuperación de información.

c. Los Links Externos

Dado que los motores de búsqueda usan medidas de popularidad de enlaces (cantidad de sitios "relacionados" o "autoritativos" que enlazan a una determinado sitio Web) para determinar los rangos de los resultados de las búsquedas, se recomienda la implementación de estrategias de "Popularidad de Enlaces", como pueden ser el alta en Directorios o el intercambio de links.

4.4 SEO de Sombreo Negro (*Black Hat SEO*)

También debe especificarse la existencia de algunas técnicas no efectivas pero que, sin embargo en la mayoría de casos, se acude por falta de información.

En la mayoría de los casos, el uso de estas técnicas no efectivas se debe a las agencias de diseño Web, que se orientan más a la visibilidad o a lo atractivo de un portal, y no a la optimización el portal para los Motores de Búsqueda. Entre este grupo de técnicas encontraremos las siguientes:

- Registro masivo en miles de buscadores
- Registro frecuente en buscadores
- Granjas de enlaces (*link faros*), suelen ser creadas por empresas para vender enlaces desde sus sitios Web.
- Empleo indiscriminado de meta-tags
- Prácticas no éticas, como el uso de texto oculto
- Páginas Web de acceso (*doorway pages*)
- Camuflaje de contenido (*cloaking*)

- Repetición de palabras sin sentido
- Rellenado de palabras claves (*keyword stuffing*)
- Texto alternativo con spam
- Otras técnicas de engaño

Además, todas estas técnicas han sido concebidas erróneamente para lograr posicionamiento de portales Web en buscadores a un menor costo sin tener que invertir en conocimiento experto acerca de posicionamiento Web en Motores de Búsqueda.

Estas pseudo-soluciones, aunque por momentos resulten efectivas, están prohibidas por los motores de búsqueda, quienes aplican severas penalizaciones sobre los portales Web que las usan, llegando a eliminar inclusive esos dominios de sus bases de datos para no presentarlos ya jamás dentro de los SERPs.

4.5 Servicios de Empresas Privadas

Diversas empresas proporcionan servicios de Posicionamiento Web. Éstas, casi en su totalidad son empresas extranjeras, que ofrecen desarrollar o modificar un portal Web para mejorar su posicionamiento en los principales Motores de Búsqueda.

Además las diversas técnicas que usan son cajas negras, es decir cada empresa mantiene en secreto su metodología, el cual va alejando del tema a la empresa contratante, cuando la idea es que las organizaciones, dueñas del portal, deben de conocer el tema ya que como conocedores de su negocio, son los mejores en administrar su información.

4.6 Cuadro Comparativo entre las Técnicas Existentes

Veamos una comparación entre las principales técnicas existentes, en la tabla 4.1:

	PPC	Técnicas SEO (SEO de Sombrero Blanco)	SEO de Sombrero Negro	Contratar Empresa
Costo	Depende del número de palabras claves que se desea manejar. Aproximadamente son \$100 diarios (25 clicks).	El costo es mínimo, implica aprender las estrategias SEO y aplicarlas en su propio portal Web.	El costo es mínimo, implica aprender las estrategias SEO y aplicarlas dentro del portal Web.	Depende de la empresa y el tipo de contrato, aproximadamente son \$4000 mensuales.
Tiempo de Implementación	días	meses	meses	meses
Beneficio	Se tiene una mayor probabilidad de conseguir tráfico.	Se logra páginas Web amigables al usuario y al motor de búsqueda, por ende un buen posicionamiento Web. El resultado es más durable. No es necesario tener que pagar mensualmente por meses o años.	Al comienzo se logra un buen posicionamiento, pero en realidad te causa daño, ya que los Motores de Búsqueda te penalizan.	Lograr un buen posicionamiento, si hace un trabajo serio.

Tiempo en Ver Resultados	En cuestión de días se verá resultados.	Los resultados se verán de 4 a 6 meses.	Los resultados se verán de 2 a 3 meses.	Es dependiente de las técnicas que usan, generalmente es cuestión de meses.
Aplica Gestión de Contenidos Web	NO	SI	NO	NO

Tabla 4.1: Comparación entre las Principales Técnicas Existentes

Capítulo 5: Propuesta

Los Motores de búsqueda, son medios de estrategias de marketing, por eso es importante que un portal Web este bien posicionado en ellos, para así conseguir clientes potenciales. Además, ubicarse en las primeras posiciones forma parte de los pilares fundamentales para el desarrollo y promoción de un Portal Web en Internet.

En este capítulo, se propone la integración de Técnicas SEO y WebCMS a fin de obtener una técnica más completa que permita promocionar un portal Web de manera más óptima. Es decir, una técnica que busca construir portales Web amigables tanto a los motores de búsqueda como a los usuarios. En la sección 5.1 se muestran los motivos que llevaron a definir la técnica propuesta. En la sección 5.2 se describe y enumera las ventajas de dicha técnica. En la sección 5.3 se enumeran las funcionalidades del prototipo. En la sección 5.4 se presenta el diagrama de Casos de uso a manera de descripción de la técnica. Y finalmente, en la sección 5.5 se mueve los diagramas de clases.

5.1 Motivación

El posicionamiento en motores de búsqueda no sólo debe buscar que el usuario final llegue al Portal Web, sino que una vez localizado, se pueda presentar de mejor forma la información. Es por eso que, estar en las primeras posiciones y mostrar la información útil al usuario, son complementarios para lograr posicionarse en las primeras SERPs.

Actualmente existen, por un lado, técnicas que buscan mejorar el posicionamiento Web pero obvian la información útil e inclusive las normas del buen diseño. Y por otro lado, existen Sistemas de Gestión de Contenidos Web (*Web Content Management System - WCMS*), que permiten manejar información y Contenidos, pero no están orientados a conseguir un buen posicionamiento Web.

A pesar de sus aplicaciones independientes, guardan estrecha relación, ya que el manejo de información por parte del WebCMS no sólo permite al portal ser útil para el usuario, sino podrá ser de gran utilidad a los motores de búsqueda para la clasificación del contenido Web, esto se debe por 2 razones: [18]

- Los motores de búsqueda son originalmente herramientas de investigación académica diseñada para encontrar información de texto, es decir, contenido.
- Los motores de búsqueda necesitan algo en el que basar sus sentencias. Cuando se escribe un término de búsqueda en un motor de búsqueda, éste ubicará algún portal con la palabra proporcionada. Por lo tanto, un sitio Web construido con pocas palabras se encuentra en una situación de desventaja desde el principio.

Otra de las motivaciones de la propuesta de este trabajo de investigación, es la existencia de algunos problemas en las técnicas actualmente utilizada para el posicionamiento Web, de las cuales destacan:

- Las técnicas existentes se centran en posicionar los Portales Web en los buscadores; más no en la accesibilidad y usabilidad del Portal (calidad de la información).
- La aplicación de recomendaciones existentes para el posicionamiento Web, conocido como SEO, queda al libre criterio de quien lo quiere aplicar, haciéndolo dependiente del mayor conocimiento de estos.
- La mayoría de las técnicas implican gastos mensuales, que para el caso de pequeñas empresas significan grandes cantidades de dinero.

En resumen, la calidad del Contenido Web también cuenta para el posicionamiento en motores de búsqueda; es por eso que, para poder conseguir un mejor posicionamiento en los Motores de Búsqueda, es necesario que los portales Web manejen normas del buen diseño e información útil así como manejar técnicas de posicionamiento Web e información indexable.

Finalmente, las empresas buscan poseer portales Web accesibles y útiles al Usuario, con información actualizada. De nada sirve estar primeros si no se puede mostrar información útil al usuario, tampoco sirve tener una página Web atractiva y bien diseñada, si no está posicionada en los principales Motores de Búsqueda. Todo se

complementa: manejar información útil, tener un diseño amigable al usuario y ser amigable a los motores de búsqueda.

5.2 Integración de Técnicas SEO y WebCMS

Como se mencionó en la sección 5.1, la mayoría de técnicas para el posicionamiento en los Motores de Búsqueda no consideran la gestión del contenido en los portales Web, muchos de ellos centran sus objetivos en ubicarse primeros, dejando de lado la importancia del diseño de las páginas y de la información que hay en ellas.

SEO y WebCMS son dos técnicas actualmente utilizadas para la optimización de portales Web pero con diferentes objetivos, la primera busca la optimización de las páginas del Portal Web, con el fin de poder posicionarlo en los principales buscadores y la segunda gestiona el contenido Web de dichos portales.

Es así que, a pesar que cada uno busca, por su lado, lograr sus objetivos; el éxito de posicionamiento de un portal Web, depende, en cierta medida, de la aplicación de ambas técnicas, es decir se requiere gestionar contenido Web y aplicar SEO para conseguir un mejor posicionamiento. En conjunto permitirán a los portales Web: 1) Proporcionar mayor información a los Motores de Búsqueda para conseguir un mejor posicionamiento, 2) Poseer páginas basadas en normas del buen diseño, 3) Proporcionar información apropiada y útil a los usuarios. 4) poseer páginas optimizadas para los motores de búsqueda al aplicar las recomendaciones SEO. En la figura 5.1 se muestra los objetivos del WebCMS y de SEO trabajar en conjunto.

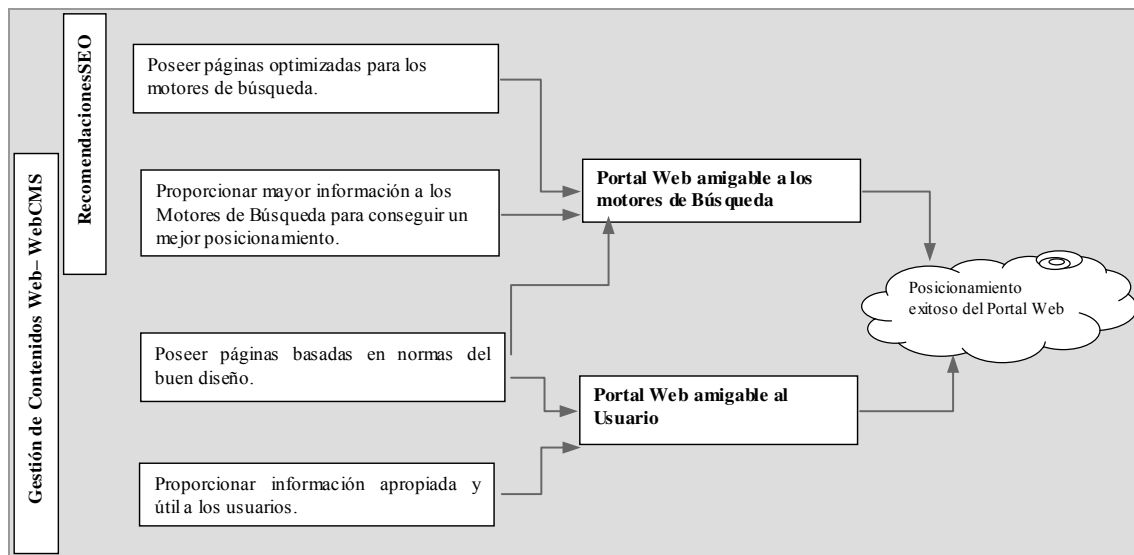


Figura 5.1: Aplicación de la Gestión de contenidos Web y de SEO para obtener un posicionamiento Web exitoso

Por todo lo expuesto anteriormente, se propone la integración de estas dos técnicas, en busca de una mejor utilidad y mejor posicionamiento Web del portal. Así se podrá contar con portales accesibles, útiles al Usuario y con información actualizada. En la figura 5.2 se puede observar el modelo conceptual de la Integración de Técnicas SEO y WebCMS que propone este trabajo de investigación.

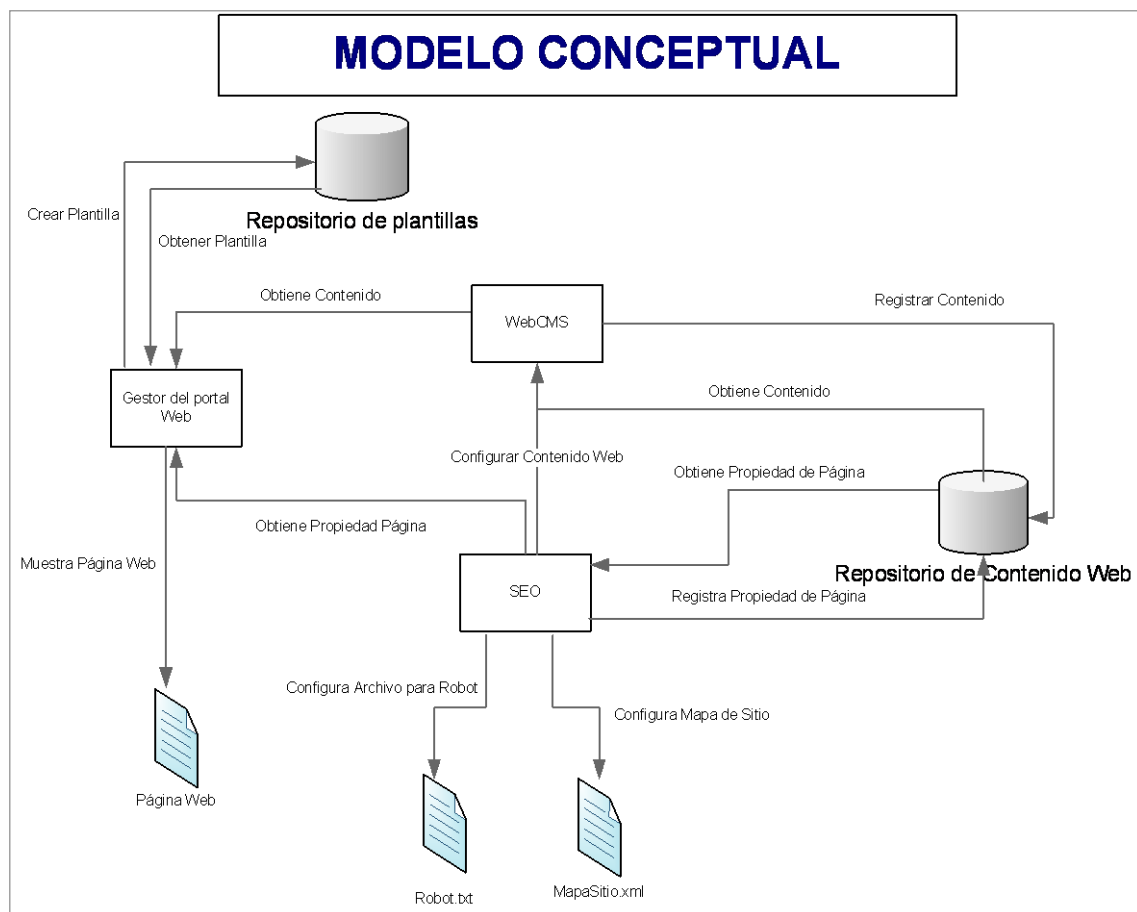


Figura 5.2: Modelo Conceptual de la Integración de técnicas SEO y WebCMS

Además, la propuesta no solo permite gestionar Contenidos Web, sino también gestionar las diferentes propiedades de la página Web, que permitirá cumplir con las más óptimas recomendaciones (técnicas) de SEO para finalmente lograr la optimización en los motores de búsqueda.

La técnica permitirá crear elementos del portal Web como por ejemplo las páginas Web a partir de plantillas ya existentes, las cuales estarán definidas bajo una estructura de diseño SEO, es decir siguiendo normas de la adecuada estructuración de páginas.

En estas páginas, el componente SEO gestionará sus propiedades, así como los archivos de configuración necesarios para los motores de búsqueda (Robot.txt, MapaSitio.xml). Y el componente WebCMS gestionará su contenido Web.

A continuación se describe con más detalle las tareas o funcionalidades de cada componente:

5.2.1 Componente SEO

Es el componente que controla y gestiona todos los aspectos del portal Web que permitan su construcción optimizado para los Motores de Búsqueda, así como el manejo de información apropiada al usuario y a los Motores de Búsqueda.

A continuación se listan las principales funcionalidades del componente SEO.

- **Gestiona Propiedades de las Páginas Web.**

Las tareas del componente SEO para la gestión de las propiedades son:

- **Controlar el mayor uso de texto y no gráficos**

La técnica fomenta el uso mayor de texto, evitando la utilización de gráficos para colocar texto en páginas. En la figura 5.3, se puede ver una página que parece tener mucho texto pero ninguno es realmente texto, cada palabra está en un archivo de imagen. Este tipo de páginas no permiten optimizar los portales Web ya que los motores de búsqueda no leen el texto en las imágenes que atraviesan, por lo que esta página no proporciona el texto que se pueden indexar por los motores de búsqueda.

Con el mayor uso de textos, se busca mayor usabilidad y evitar el mal diseño, ya que todas las imágenes tomar mucho más tiempo para descargar el texto equivalente que tendría.

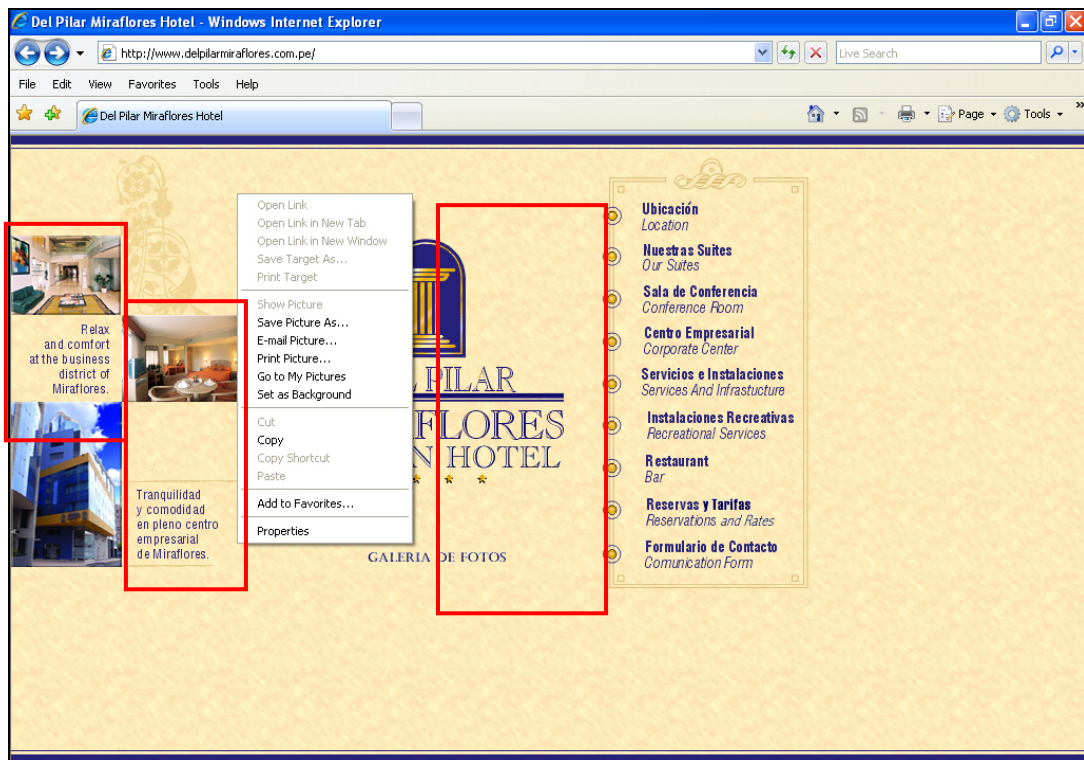


Figura 5.3: Uso de imágenes y no texto en el Portal Web del hotel Del Pilar Miraflores

- **Permitir la configuración de las propiedades de las imágenes**

La técnica permite el uso de imágenes con atributos de texto, el componente SEO gestiona y configura las propiedades de los elementos de las páginas a fin de que cada uno de estos, cumpla las recomendaciones SEO.

El componente SEO permite incluir atributos 'ALT TEXT' en las imágenes de una página Web, para así proporcionar una descripción a cada una de ellas.

La importancia del atributo 'ALT TEXT', es debido a que los motores de búsqueda utilizan el texto para tener más idea del tema de la imagen. La técnica fomenta el uso de las palabras clave en el atributo ALT.

- **Permitir el uso de enlaces con textos largos**

El componente SEO permite crear enlaces que puedan explicar a que se hace referencia. Además permite, mediante el uso de palabras clave en los enlaces, decir a los motores de búsqueda de que tema se refiere la página Web indexada.

- **Corregir Ortografía y edición**

El componente SEO controlará la edición de páginas y corrige la edición de errores. Las páginas sin errores no sólo hacen que el portal Web parezca más profesional a los usuarios visitantes, sino permiten asegurarse que la información que se muestra no es en vano.

- **Optimizar el lugar donde aparecen las palabras claves**

El componente SEO permite optimizar la ubicación de las palabras claves del portal Web, es decir, ubicar estas palabras en las etiquetas fundamentales de una estructura de página, como son: TITLE, META DESCRIPTION, META ROBOT, CABECERAS H1, H2 Y H3.

- **Genera Plantillas para las Páginas Web**

Se utiliza plantillas que contendrán la estructura básica de las páginas Web del portal. Estas plantillas, como recomienda SEO, tendrán un diseño simple y limpio, por ejemplo usar texto negro y fondo blanco.

- **Gestiona Navegabilidad del Portal Web**

El componente SEO permite gestionar diferentes elementos para facilitar la navegación del Portal Web.

La técnica evita hacer difícil a los usuarios visitantes el desplazarse a través de un sitio Web. Permite que los usuarios encuentren rápidamente lo que necesitan.

La gente piensa diferente, por lo que necesita proveerlo con numerosas formas para encontrar su camino en el portal Web. Y al hacerlo, se está dando más información a los motores de búsqueda y la garantía de que los motores de búsqueda puede navegar su sitio fácilmente.

Para esta tarea, el módulo SEO se encargará de construir y gestionar los archivos de navegación como son los sitemap.xml.

Además, se encargará de crear y configurar páginas con enlaces a las diferentes áreas del portal Web, por ejemplo una página con Tabla de Contenidos o Índice de la página, Barras de navegación y Enlaces textuales de Navegación.

- **Gestiona Contenido Web**

El componente SEO busca manejar información útil en el portal Web por lo cual permite la edición de textos, es decir la edición del contenido de la página Web.

Para esta tarea dispone de un componente WEBCMS que se centrará en la gestión de la información del portal.

5.2.2 Componente WebCMS

El prototipo, mediante el módulo de WebCMS, permite crear una estructura de soporte para la creación y administración de contenidos Web.

Permite controlar una o varias bases de datos donde se aloja el contenido del sitio. Además, permitirá manejar de manera independiente el contenido y el diseño. Así, es posible manejar el contenido y darle en cualquier momento un diseño distinto al sitio sin tener que darle formato al contenido de nuevo, además de permitir la fácil y controlada publicación en el sitio a varios editores.

5.2.3 Componente para la Gestión del Portal Web

El prototipo consta de un Módulo que permite configurar las plantillas en las que se basará las páginas del portal. Además permite la creación, eliminación, modificación de páginas y demás elementos del Portal Web.

Las tareas de este componente se muestran a continuación:

- **Limita el Uso de archivos Multimedia**

El prototipo limita el uso de archivos multimedia. Las plantillas de las que se dispone no permiten el uso excesivo de estos archivos, principalmente archivos flash. Adicionalmente se cuenta con un servicio de ayuda que proporciona información al usuario acerca de las recomendaciones SEO.

Muchas veces resulta contradictorio el uso de archivos flash, ya que en ocasiones los diseñadores ven en lo elementos multimedia una forma de darle visibilidad al diseño, pero más que hacerlo llamativo, puede estar construyéndose páginas muy poco amigables a los motores de búsqueda e inclusive a los Usuarios.

En la figura 5.4 se muestra un ejemplo, la página principal de un portal Web peruano (<http://www.hikari.com.pe>), el cual está diseñado enteramente en flash. Todos los textos e imágenes están inmersos en este tipo de archivo.

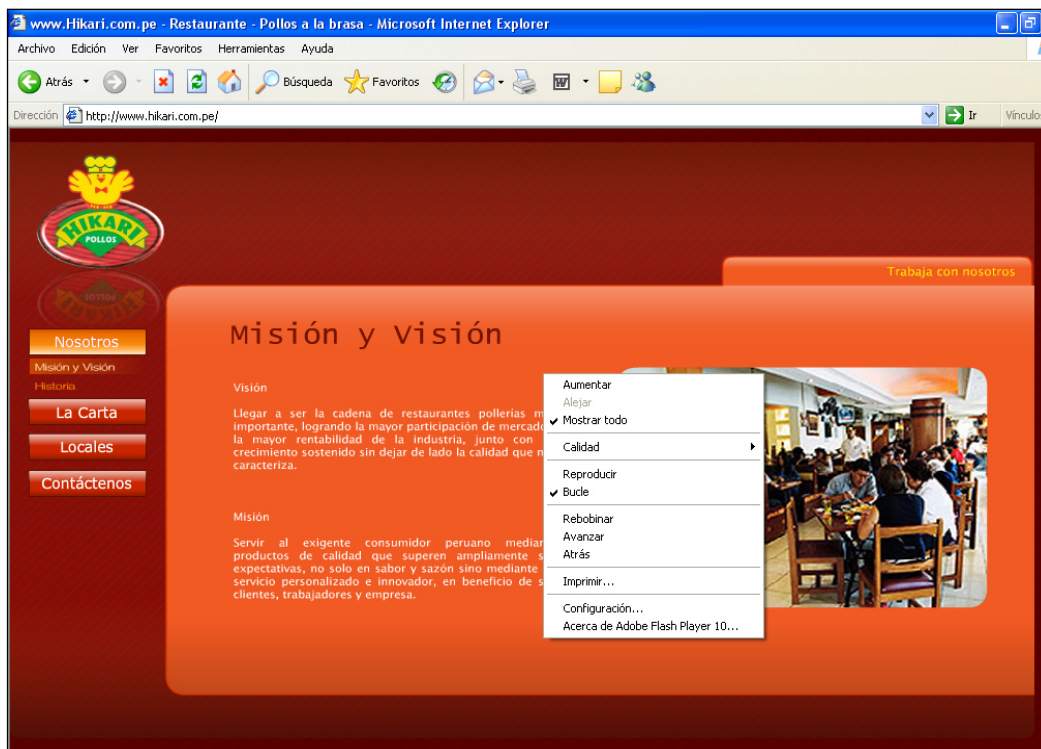


Figura 5.4: Diseño del portal Web del Restaurante Hikari basado totalmente en flash

Así en la figura 5.5 se muestra como Google registra en la caché a este portal Web, y se ve que no muestra ningún texto al que pueda indexar.

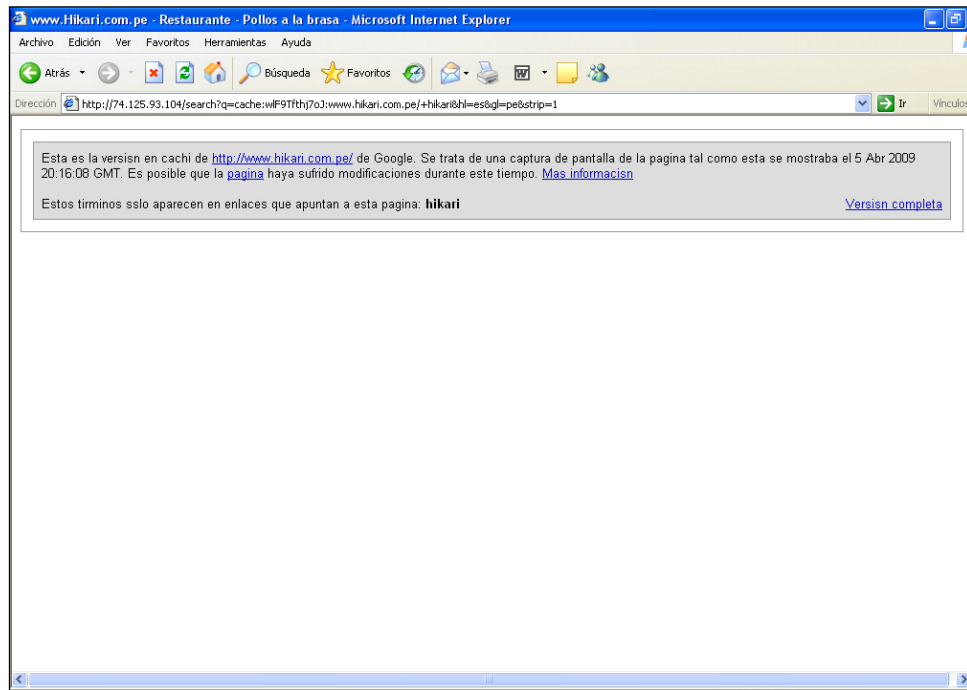


Figura 5.5: Registro en caché versión sólo texto

Es resumen, la técnica propuesta plantea tomar las mejores técnicas SEO, así como los conceptos de Gestión de Contenidos a fin de reunir las mejores prácticas de optimización para Motores de Búsqueda, logrando una herramienta más eficiente.

Además, permite al propietario, conocedor de su negocio, definir y/o actualizar diferentes propiedades de la página Web, así como el contenido de cada una de ellas.

Los principales beneficios que se pueden obtener de la utilización de la integración de SEO y WebCMS se pueden enumerar de la siguiente manera:

- Busca mejorar el posicionamiento porque no sólo se centra en el contenido de la página, sino también en las propiedades del portal Web.
- Permitirá definir la estructura de las páginas a fin de ser amigable a los motores de búsqueda.
- Permitirá construir las páginas Web a fin de ser amigable a los Usuarios.
- Es una herramienta que va a ser muy útil de muchas maneras, para los motores de búsqueda y a la vez para los usuarios, actualizar contenido

- Proporcionará a los Motor de Búsqueda mayor posibilidad de encontrar palabras claves.
- Permitirá al usuario, administrador del portal y conocedor de la empresa, actualizar el contenido en tiempo real y a la vez modificar aquellas propiedades recomendadas por SEO.
- Mediante el prototipo, los usuarios accederán a las técnicas SEO más relevantes, es decir, proporcionará mayor conocimiento y accesibilidad de las técnicas SEO.
- Mejorar la etapa de mantenimiento del portal Web.
- Permitirá aplicar Normas del buen diseño, las cuales también son considerados por los motores de búsqueda.
- Inclinación por el contenido y manejo de la información
- Un buen diseño ayudarán no sólo a los visitantes del portal sino al trabajo de los motores de búsqueda sobre el portal.

5.3 Funcionalidades de la Integración Propuesta

Tomando como punto de partida las necesidades para el posicionamiento Web, las características principales de esta aplicación se pueden resumir en los siguientes puntos:

- **Accesibilidad y navegación, facilitar la indexación por parte del robot.** Es importante la accesibilidad y navegación de un sitio Web. El Sistema de Gestión de Contenidos Web (WCMS) servirá de soporte para no presentar obstáculos a los robots de búsqueda.
- **Manejo y control de Meta Tags y Títulos.** Para cada página del portal Web, el sistema permitirá la edición de los meta tags “Keywords” y “Description”, así como definir por separado el título de la información y el título de la página Web.
- **Soporte de estándares XHTML y CSS.** Tener la Web montada sobre los más nuevos estándares de forma correcta puede ayudar mucho al posicionamiento Web, tanto *XHTML* como *CSS* ofrecen ventajas al posicionamiento. El sistema garantizará que los contenidos generados y editados desde la misma cumplan con estos estándares.

- **Mapa del sitio, y Google SiteMaps.** El sistema será capaz de generar automáticamente el mapa del sitio, tanto para dar a conocer la estructura a los usuarios como para asegurar que hay al menos una vía para el robot de llegar a todos los contenidos.
- **Manejo del archivo ROBOTS.TXT.** Desde el sistema se manejará de forma transparente el archivo Robots.txt para indicar que URLs deseamos incluir en el listado de los motores de búsqueda.
- **Control de Errores 404.** El manejo de errores 404, es otra de las capacidades del prototipo. Se usa una página de error personalizada, para la optimización del portal Web para buscadores.
- **Redirects 301.** El sistema debe usar los redirects 301 para redireccionar urls con contenido antiguo, para la optimización del portal Web para buscadores.
- **Manejo de Canonizaciones.** El prototipo permite el manejo de canonización de Urls para la optimización SEO del portal Web. Es decir que el portal pueda ser reconocido por una única Url de dominio por los usuarios navegantes y por los motores de búsqueda.
- **Reportes estadísticos de accesos y uso.** Las estadísticas de accesos a información y uso de los servicios, pueden ser ofrecidas por el propio CMS mejor que muchos sistemas de estadísticas que mezclan los accesos de usuarios con robots de búsquedas y otros tipos de accesos. Por esa razón es útil disponer de cierto tipo de estadísticas dentro sistema de gestión de contenidos que complemente las obtenidas por los registros del servidor Web u otros sistemas.
- **Base de Artículos e información de posicionamiento Web.** Interesante y útil sería la posibilidad que como parte de la documentación del sistema se incluya una amplia base de conocimientos dirigida a los editores y gestores de información para propiciar el uso correcto del sistema y evitar de errores de estos que puedan ser calificados como spam o técnicas penalizables por parte de los robots de búsqueda.

No olvidemos que aunque los sistemas de gestión de contenidos deben potenciar el posicionamiento Web, aun así no se podría llevar del todo correcto si no se pone a trabajar el talento humano en dicha función.

5.4 Casos de Uso

En esta sección se especifica la forma de comunicación existente entre los usuarios y las funcionalidades de la técnica propuesta. En la figura 5.6 se muestra el diagrama de casos de uso hechos en la fase de Análisis.

Además, para una mejor comprensión, se adjunta como Anexo A, el documento de Especificación de Casos de Uso.

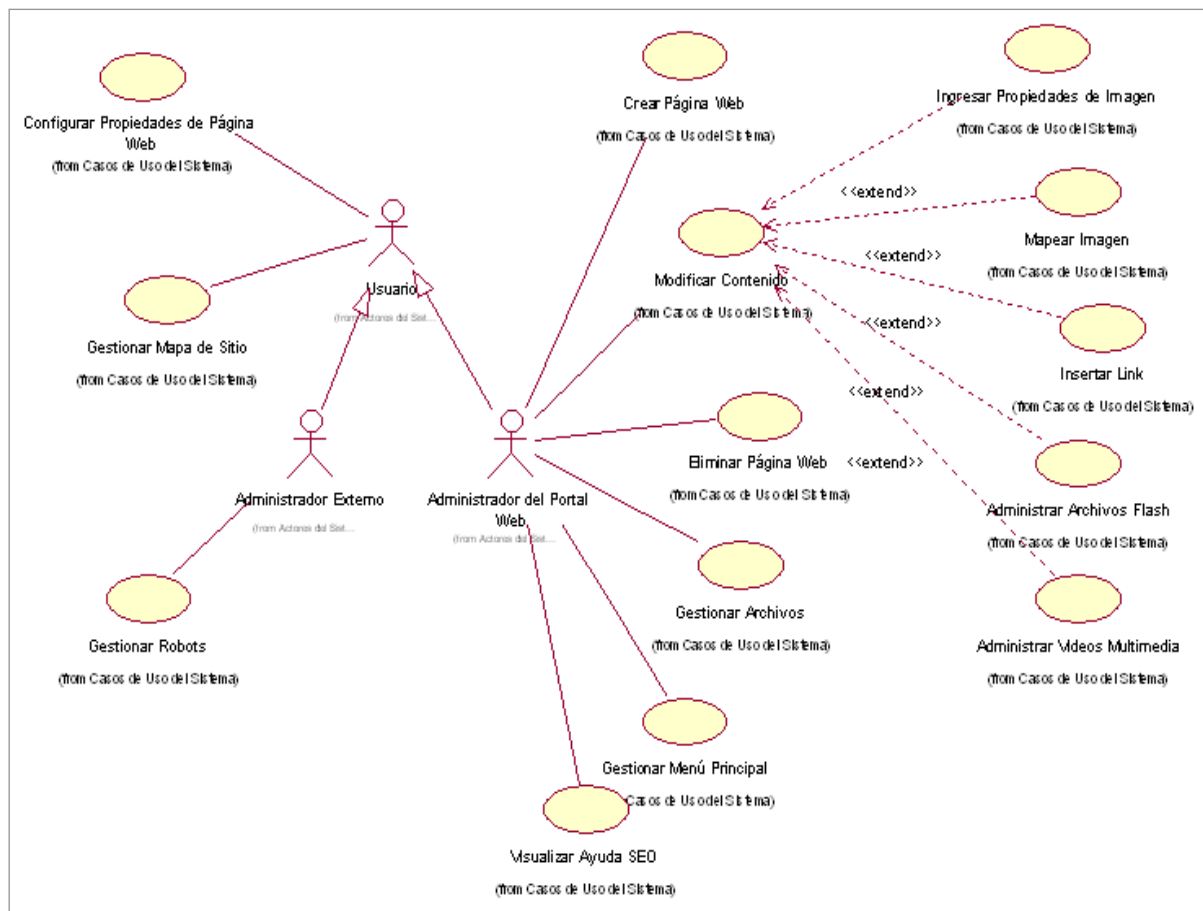


Figura 5.6: Diagrama de Casos de Uso.

5.5 Diagrama de Clases

En esta sección se describe la estructura y el modelo Conceptual de la información que se manejará en el sistema.

Se muestra el diagrama de clases en la figura 5.7, el cual ha sido hecho como parte del análisis del Prototipo y muestra el modelo Conceptual.

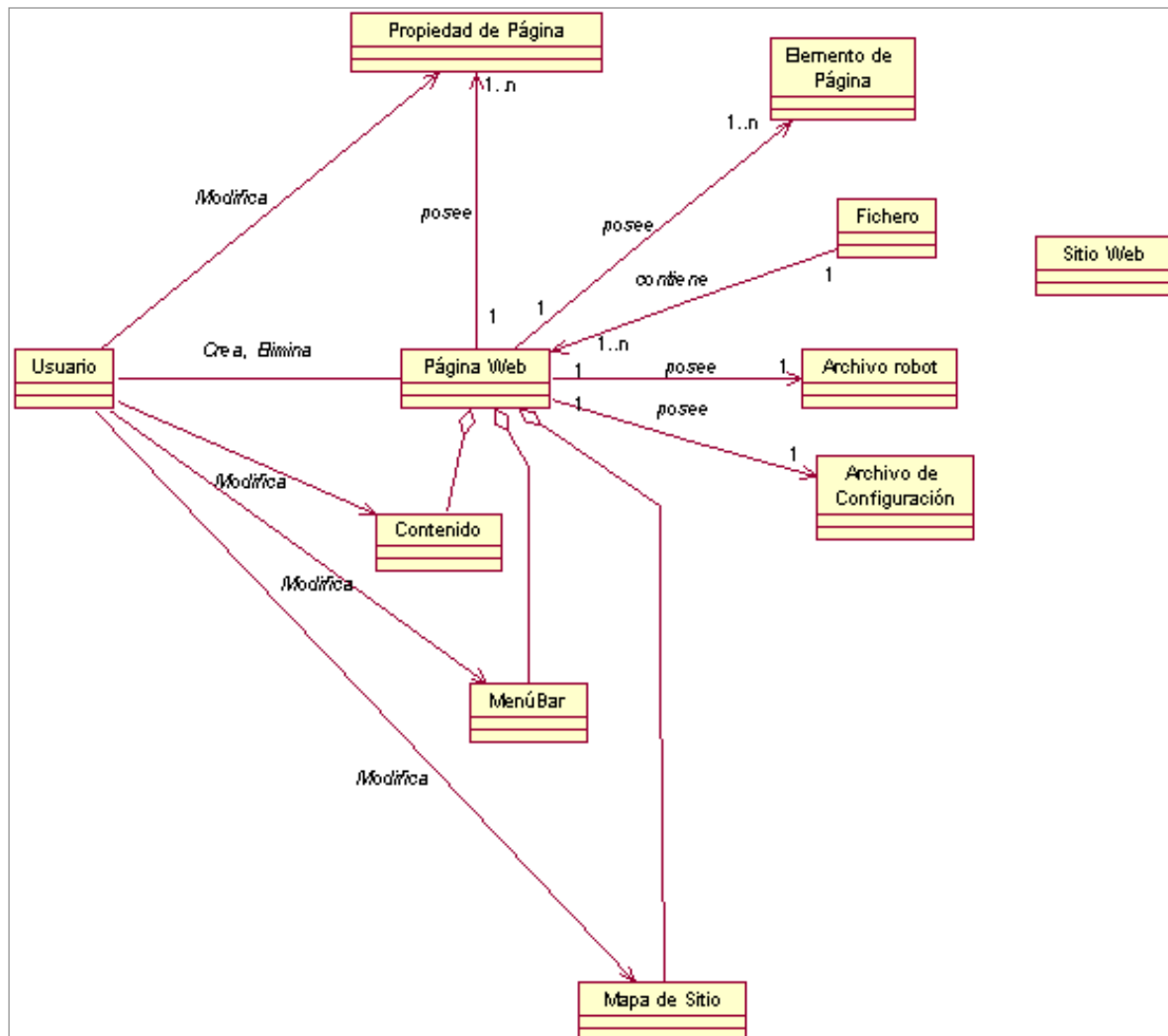


Figura 5.7: Diagrama de Clases.

Capítulo 6: Validación de la propuesta

Este capítulo describe el prototipo que fue construido, en base a la propuesta de integración de Técnicas SEO y WebCMS descrita en el capítulo 5. El prototipo a presentar permite crear páginas Web, así como la administración de propiedades de éste, con el fin de obtener portales Web optimizados, tanto para los motores de búsqueda como para los usuarios.

Actualmente, existen Sistemas Web que gestionan Contenidos, estas permiten crear páginas Web, pero no administrar ciertas propiedades importantes para la optimización del portal Web a fin de conseguir un mejor posicionamiento.

Lo que se busca es un sistema (prototipo) que permita administrar todos aquellos elementos que, en su conjunto, optimizan los portales Web, es decir un sistema que gestione toda la información del portal, manteniéndola actualizada, y además gestione toda los diferentes factores, que según las recomendaciones SEO, son de suma importancia para mejorar el posicionamiento Web.

En la sección 6.1 se describirá el prototipo implementado, en la sección 6.2 se muestra el caso de estudio elegido, así como la forma en que se aplicó la solución propuesta y los resultado respectivos. Por último, en la sección 6.3 se hace un resumen de este capítulo.

6.1 Descripción del Prototipo

En esta sección, se da a conocer el alcance de la propuesta así como los diagramas más relevantes resultado de la etapa de análisis y diseño del prototipo.

6.1.1 Alcance

El prototipo desarrollado, llamado GECOSEO, consiste en una aplicación Web que propiciará la optimización de los portales Web a fin de conseguir un mejor posicionamiento en los principales buscadores Web.

La optimización de un Portal, se consigue gracias a que el prototipo permite crear un conjunto de páginas web en base a las normas del buen diseño y a las recomendaciones SEO, esto permitirá obtener páginas optimizadas para los motores de búsqueda.

Además, el prototipo permite gestionar las principales propiedades y elementos de las páginas, basadas también en las normas del buen diseño y a las recomendaciones SEO.

Por otro lado, permite gestionar, de forma uniforme y dinámica la información mostrada en el portal.

6.1.2 Objetivos

En esta sección, se da a conocer los principales objetivos del prototipo implementado. Tales objetivos pueden resumirse en:

- Poseer páginas optimizadas para los motores de búsqueda.
- Proporcionar mayor información a los motores de búsqueda para conseguir mejor posicionamiento.
- Poseer Páginas, cuyo código html sea basadas en las normas de buen diseño
- Proporcionar información apropiada y útil los usuarios visitantes.

6.1.3 Diseño Físico

A continuación, se muestra el diseño del prototipo, véase la figura 6.1.

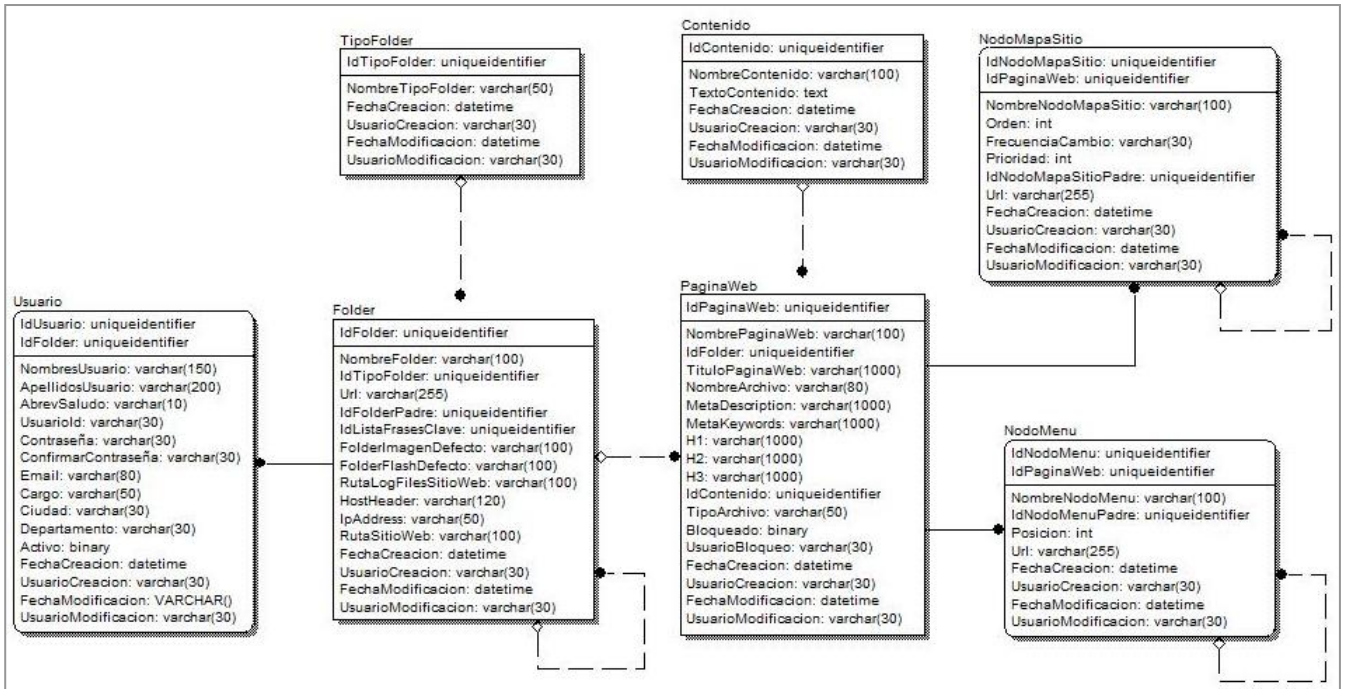


Figura 6.1: Diseño Físico del prototipo.

6.1.4 Diagramas de despliegue

Se muestra con mayor detalle los componentes, con los que cuentan el Sistema así como la relación entre ellas. Véase la figura 6.2.

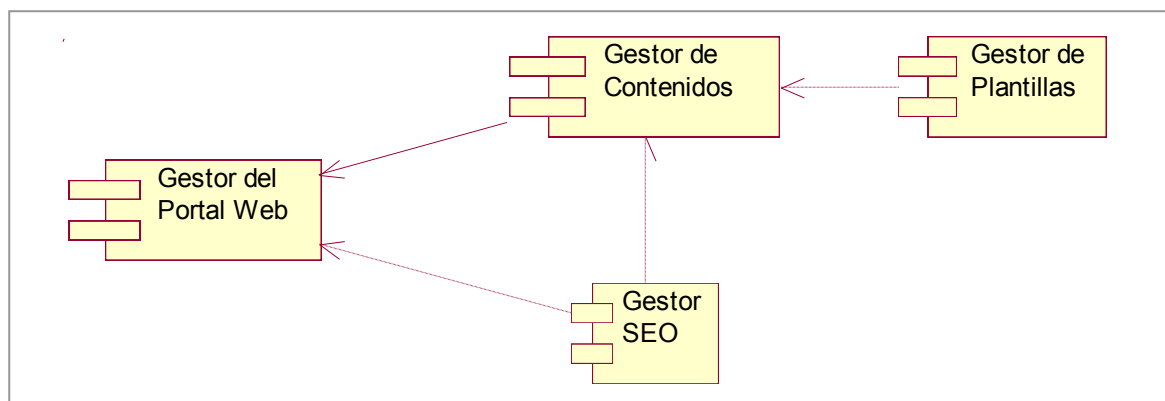


Figura 6.2: Diagrama de Despliegue.

6.2 Funcionalidad del Prototipo

En esta sección se muestran y detallan las funcionalidades del prototipo a través de sus principales interfaces de Usuario.

6.2.1 Acceso a la Administración del Contenido: Uso de Usuario y Contraseña

El prototipo permite a los usuarios administradores del portal Web, acceder al Sistema de Gestión de contenidos y SEO, mediante un usuario y una contraseña, como se muestra en la figura 6.3.



The screenshot displays the login interface for 'SOFTBRILLANCE SISTEMAS EMPRESARIALES'. The header includes the company logo, contact information (Tel: (511) 272 - 0513, info@sbperu.net), and language options (English, Español). A navigation bar lists various sections: HOME, QUIENES SOMOS, CLIENTES, SOLUCIONES, SERVICIOS, CASOS DE EXITO, BLOGS, and CONTACTENOS. On the left, a sidebar menu categorizes system modules under 'SISTEMA INTEGRADO SBA', including 'FINANZAS Y CONTABILIDAD' (with sub-items like CONTABILIDAD, CUENTAS POR PAGAR, etc.), 'LOGISTICA', 'RECURSOS HUMANOS', 'COMERCIAL', 'ACTIVOS FIJOS', 'MANTENIMIENTO', 'PRODUCCION', and 'PRESUPUESTO'. The main content area is titled 'WCMS - Iniciar Sesión' and features a login form with fields for 'Usuario:' and 'Contraseña:', followed by an 'Entrar' button.

Figura 6.3: Página de Inicio de Sesión.

6.2.2 Menú Principal:

Una vez ‘logueado’ en el Sistema, el usuario administrador dispone de un Menú Principal (véase la figura 6.4), el cual contiene las siguientes opciones:

- Modificar Página Web.
- Crear Página Web.
- Eliminar Página Web.
- Modulo de Administración.
- Salir.



Figura 6.4: Menú Principal.

- **Opción: Modificar Página Web**

Esta opción permite al Usuario Administrador elegir cualquier página del portal Web y modificar su contenido, para esto se dispone de un Editor HTML, tal como se muestra en la figura 6.5.



Figura 6.5: Página para la modificación de Contenido Web.

La modificación de las páginas Web, permite Guardar y Publicar Página Web así como Configurar Propiedades de la Página Web (véase la figura 6.6).



Figura 6.6: Subopciones para la modificación de una Página Web.

Para Modificación de páginas Web se dispone de las siguientes subopciones:

- Guardar y Publicar Página Web.
- Guardar Página Web para Aprobar.
- Configurar Propiedades de la Página Web.
- Ayuda SEO.
- Deshacer Cambios.

A continuación detallamos cada uno de ellas.

- a. **Guardar y Publicar:** Permite guardar los cambios y publicarlos para que los cambios se reflejen en tiempo real.
- b. **Guardar para Aprobar:** A diferencia del anterior, esta subopción permite guardar los cambios, pero no los publica sino los mantienen como pendiente de publicación.
- c. **Configurar Propiedades de la Página Web:** Permite configurar las propiedades más importantes de la página Web para la optimización SEO, como son: el Título de la Página Web, Meta Keyword, Meta Description, y los encabezados H1, H2 y H3, tal como se muestra en la figura 6.7.

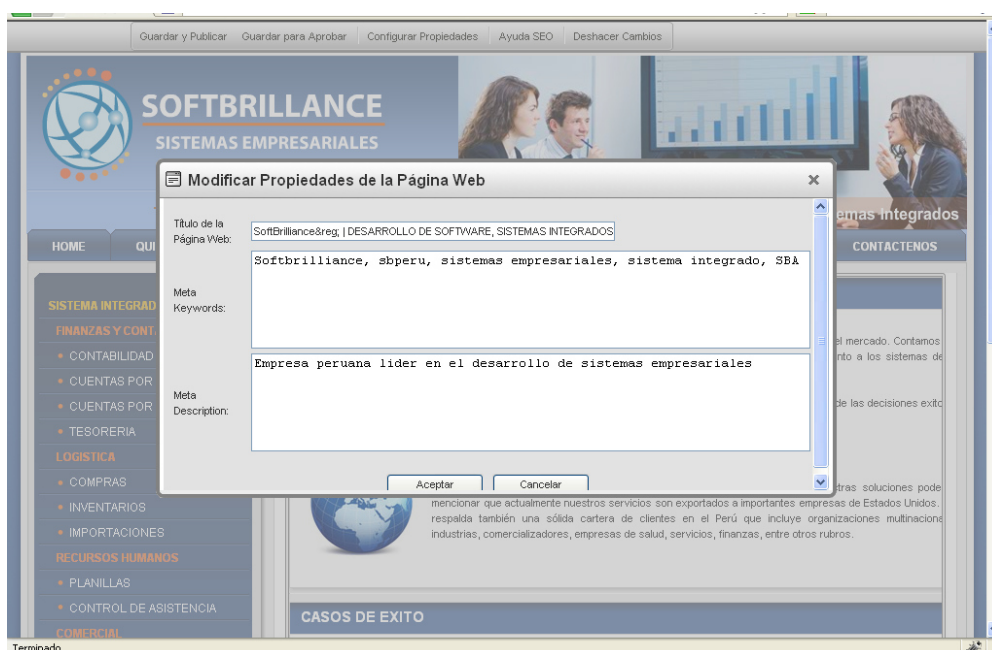


Figura 6.7: Configuración de propiedades de la página Web

d. Deshacer Cambios: Cancela los cambios que se están realizando en ese momento.

- **Editor HTML:**

Un apartado especial es para el Editor de Contenidos de Páginas Web. Este está disponible para la modificación y edición de contenidos Web, véase la figura 6.8.



Figura 6.8: Editor de HTML

El Editor HTML, posee diversas herramientas, descritas a continuación:

a. Revisión de Ortografía y Gramática:

Esta herramienta de ayuda SEO, corrige la ortografía y Gramática, proporcionando mayor utilidad a la información tanto para los usuarios como para los motores de búsqueda.

b. Portapapeles:

Esta herramienta pone a disposición, diversas funciones que facilitan la edición de contenidos, como pro ejemplo:

- Cortar Texto.
- Copiar Texto.
- Pegar.
- Pegar desde Word-
- Pegar desde Word, quitando Fuente.
- Pegar Texto Plano.
- Pegar como Html.

c. Manejo de fuentes de textos:

Pone a disposición, diversas funciones que facilitan la edición de contenidos. Esta opción permite:

- Negrita.
- Cursiva.
- Subrayado.
- Color de Fuente.
- Color de Resaltado del Texto.
- Cambiar el Tipo de Fuente.
- Cambiar el Tamaño de la Fuente.
- Seleccionar Todo.
- Deshacer.
- Repetir.

d. Manejo de Párrafo de textos:

Pone a disposición, diversas funciones que facilitan la edición de contenidos. Esta opción permite:

- Alinear Texto a la Izquierda.
- Centrar.
- Alinear Texto a la Derecha.
- Justificar.
- Nuevo Párrafo.
- Aplicar Sangría a la Izquierda.
- Aplicar Sangría a la Derecha.

- Viñetas.
- Numeración.
- Insertar Línea Horizontal <HR>.

e. Tablas:

Permite el manejo de tablas en la página Web, tal como se muestra en la figura 6.9. Esta opción permite:

- Insertar Tabla.

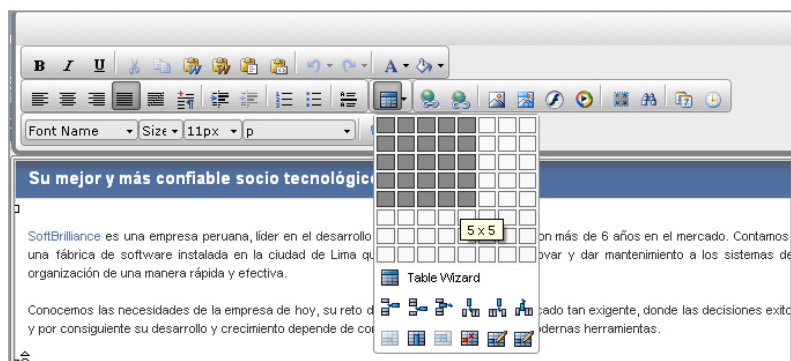


Figura 6.9: Editor de HTML – tablas.

f. Edición de Enlaces Internos y Externos:

Esta opción permitirá gestionar los enlaces internos y/o externos de la página Web (ver la figura 6.10). Permitirá agregar enlaces optimizados de acuerdo a SEO (ver la figura 6.11). Esta opción permite:

- Insertar Hipervínculo.
- Remover Hipervínculo.

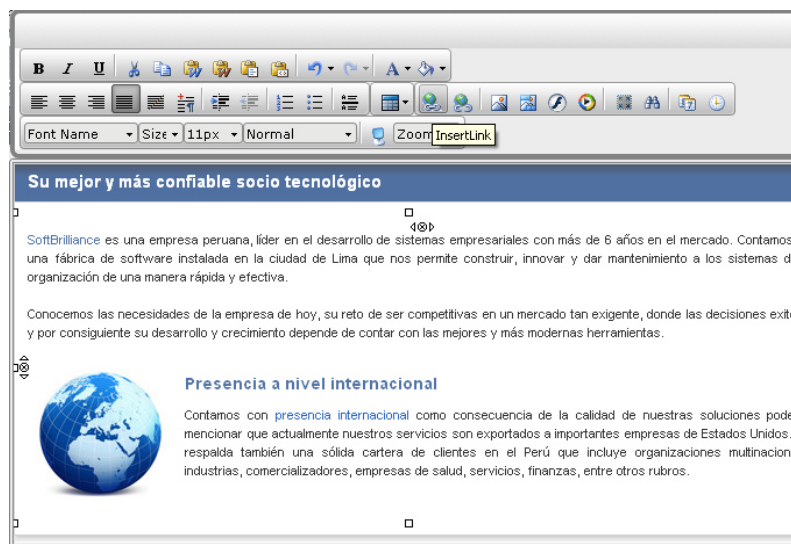


Figura 6.10: Opción de manejar enlaces en el Editor de HTML

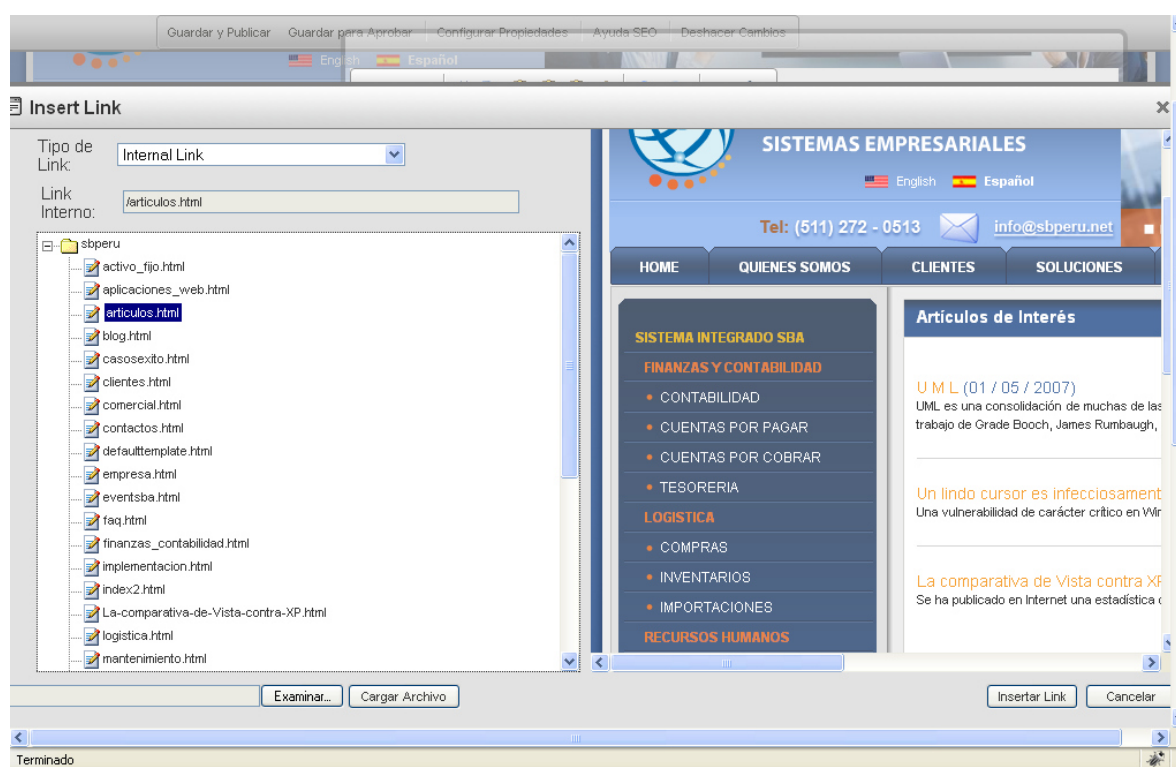


Figura 6.11: Seleccionar la página interna o dirección externa a la que se enlazará.

g. Imágenes:

Esta opción permitirá gestionar las imágenes optimizados de acuerdo a SEO, tal como se ve en la figura 6.12. La gestión de imágenes incluye:

➤ Insertar Imagen.

Las imágenes insertadas serán optimizadas según SEO y formará parte del contenido de la página Web, tal como se muestra en la figura 6.13.

➤ **Modificar Propiedades de la Imagen.**

Modificar las propiedades de una imagen, es otra manera de poseer imágenes optimizadas según SEO, véase la figura 6.14.

➤ **Mapear Imagen.**

Permite manejar mapeo de imágenes para una mayor optimización según SEO como parte del contenido de la página Web, tal como se muestra en la figura 6.15.



Figura 6.12: Opciones para la gestión de imágenes.

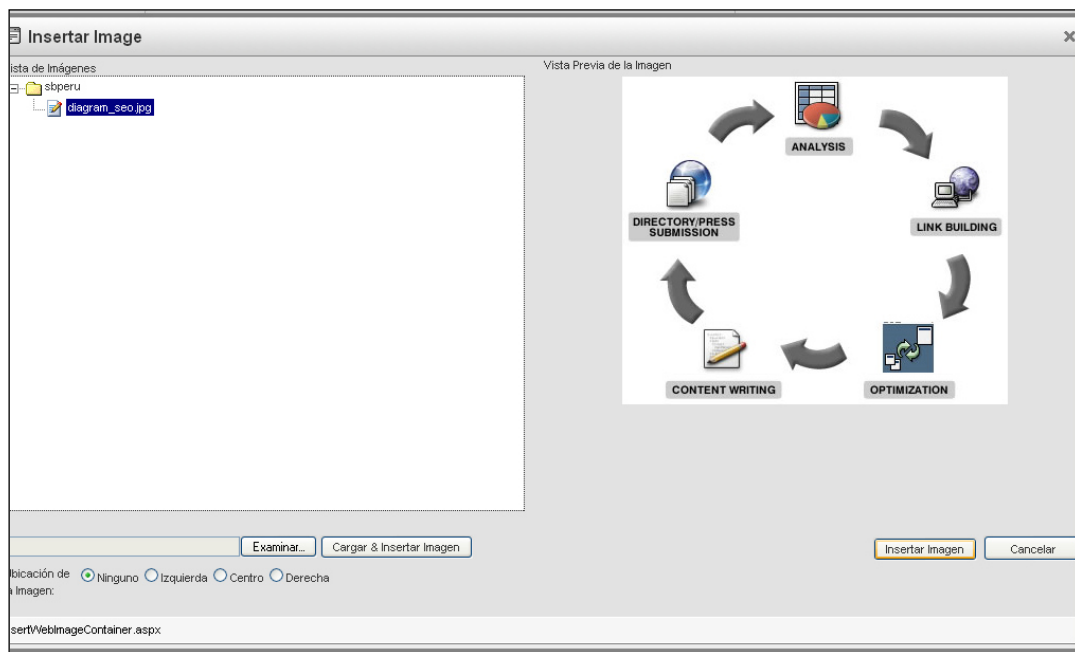


Figura 6.13: Inserción de imágenes.

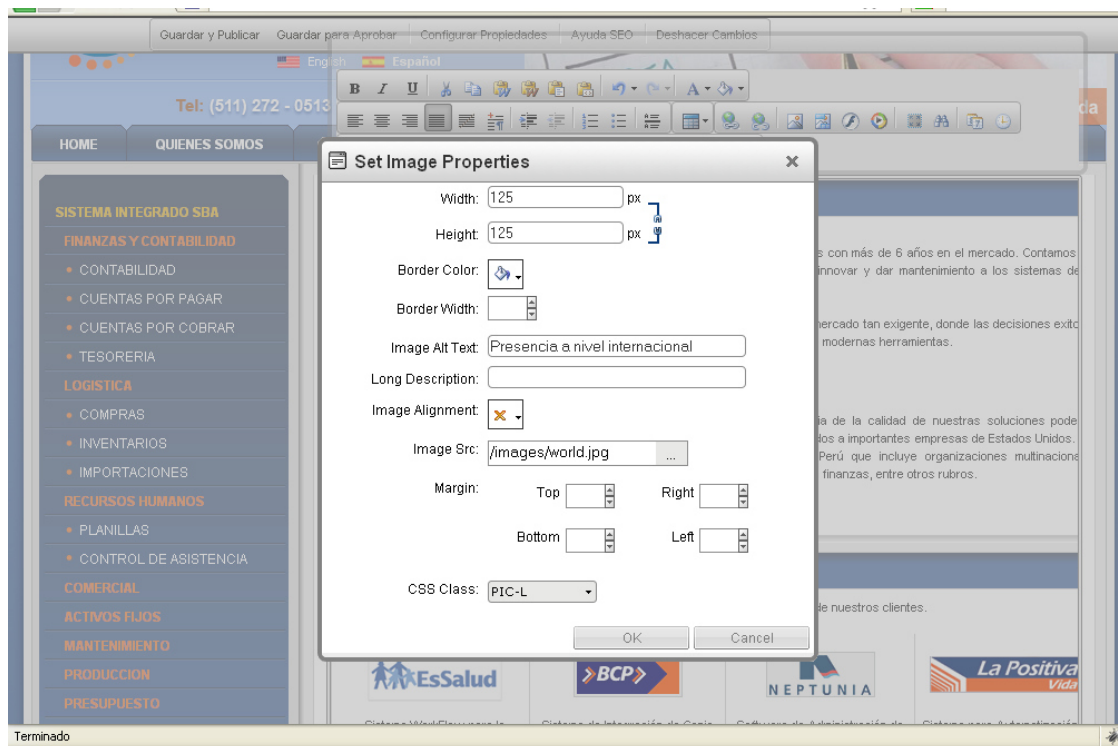


Figura 6.14: Configurar las propiedades de la imagen.

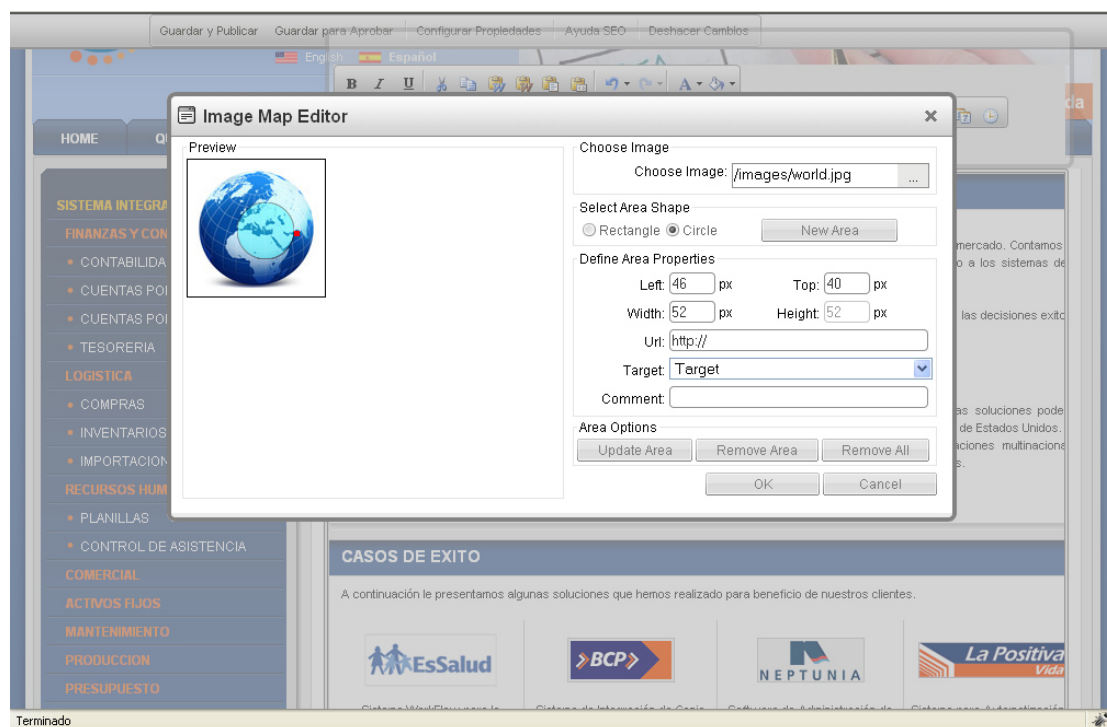


Figura 6.15: Editor de 'Image Map'.

h. Objetos:

Esta opción permite la gestión de otros elementos que enriquecerán el contenido de una página Web (ver figura 6.16). La gestión de los objetos incluye:

- Administrador de Flash.
- Administrador de Videos Multimedia.
- Insertar Fecha.
- Insertar Hora.

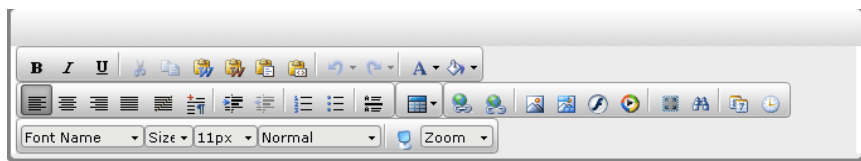


Figura 6.16: Opciones para la gestión de objetos multimedia.

● Opción: Crear Página Web

Esta opción le permite al usuario administrador, crear una nueva página Web para el portal, ya sea desde una página en blanco o una copia de alguna ya existente, como se muestra en la figura 6.17.

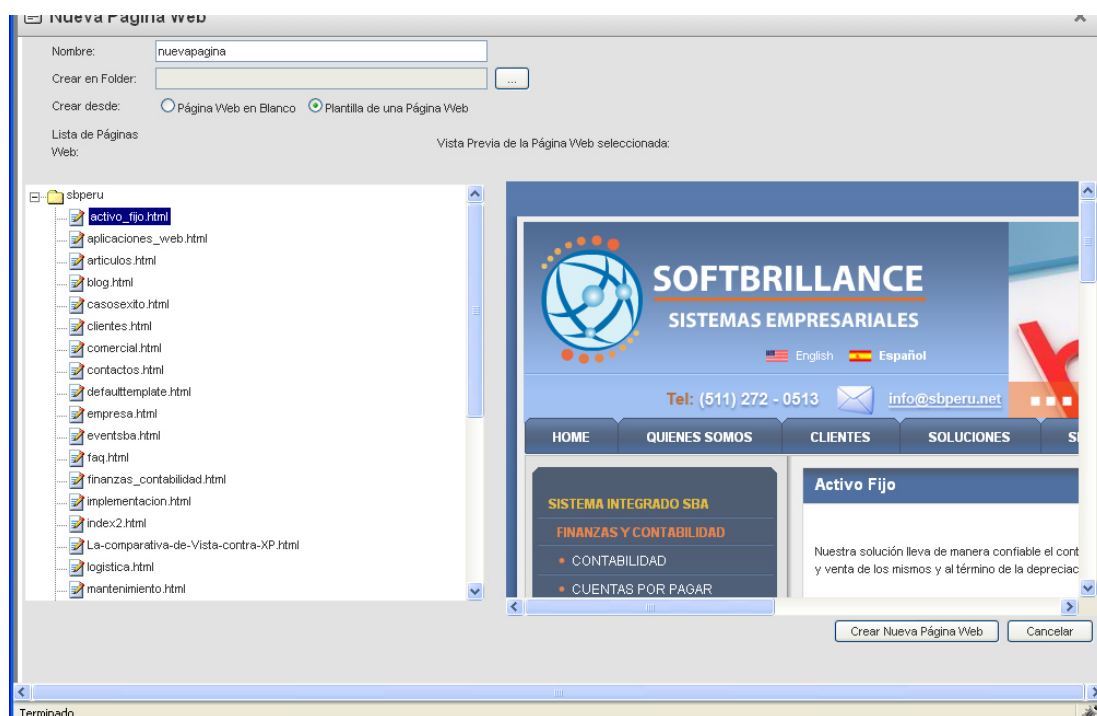


Figura 6.17: Opción para la creación de páginas Web.

- **Opción: Eliminar esta Página Web**

Esta opción permite al Usuario Administrador eliminar determinada página Web como parte del Mantenimiento del portal Web, tal como se muestra en la figura 6.18.

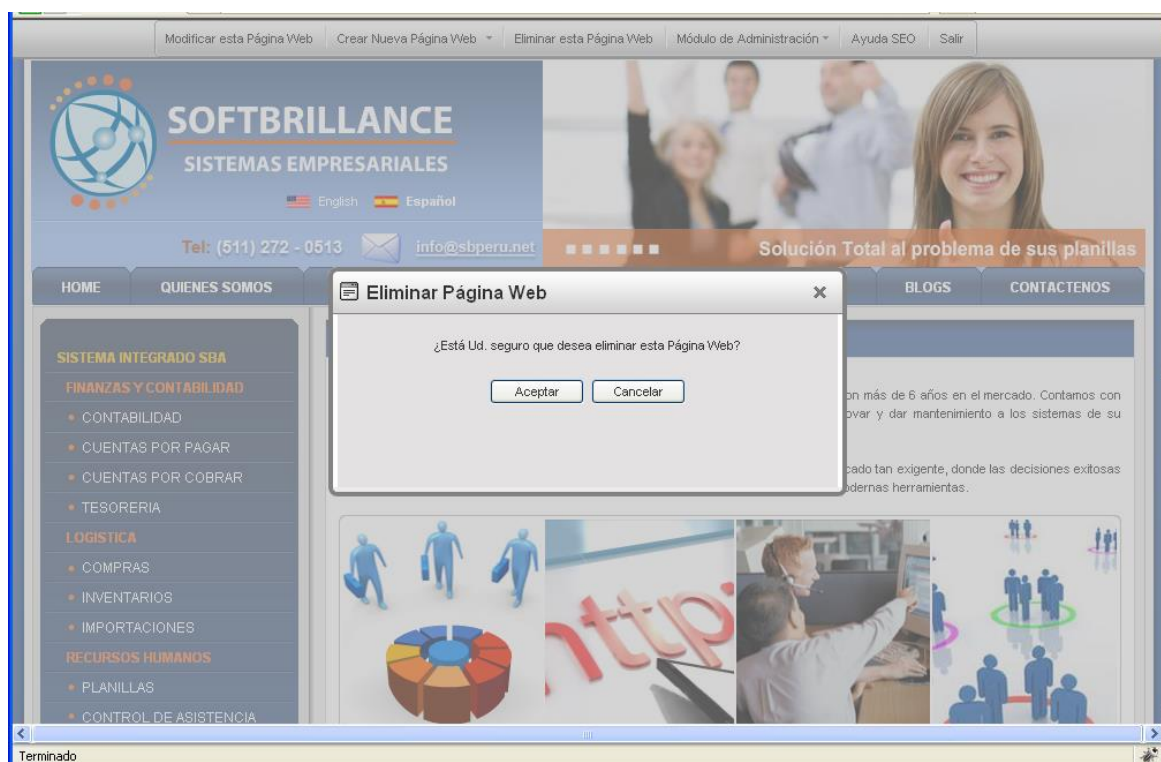


Figura 6.18: Opción para la eliminación de una página Web.

- **Opción: Administración del Portal Web**

El Módulo de Administración permite gestionar diferentes elementos importantes del portal Web, a fin de configurar las características y/o propiedades del portal para la optimización SEO, tal como se muestra en la figura 6.19. Las subopciones con las que cuenta son las siguientes:

- Administración de Archivos.
- Administración del Mapa de Sitio.
- Administración del Menú Principal.
- Administración del archivo robot.txt



Figura 6.19: Opciones del Módulo de Administración.

a. Administración de Archivos

Permite gestionar archivos en el Portal Web, es decir cargar o eliminarlos, ver la figura 6.20.

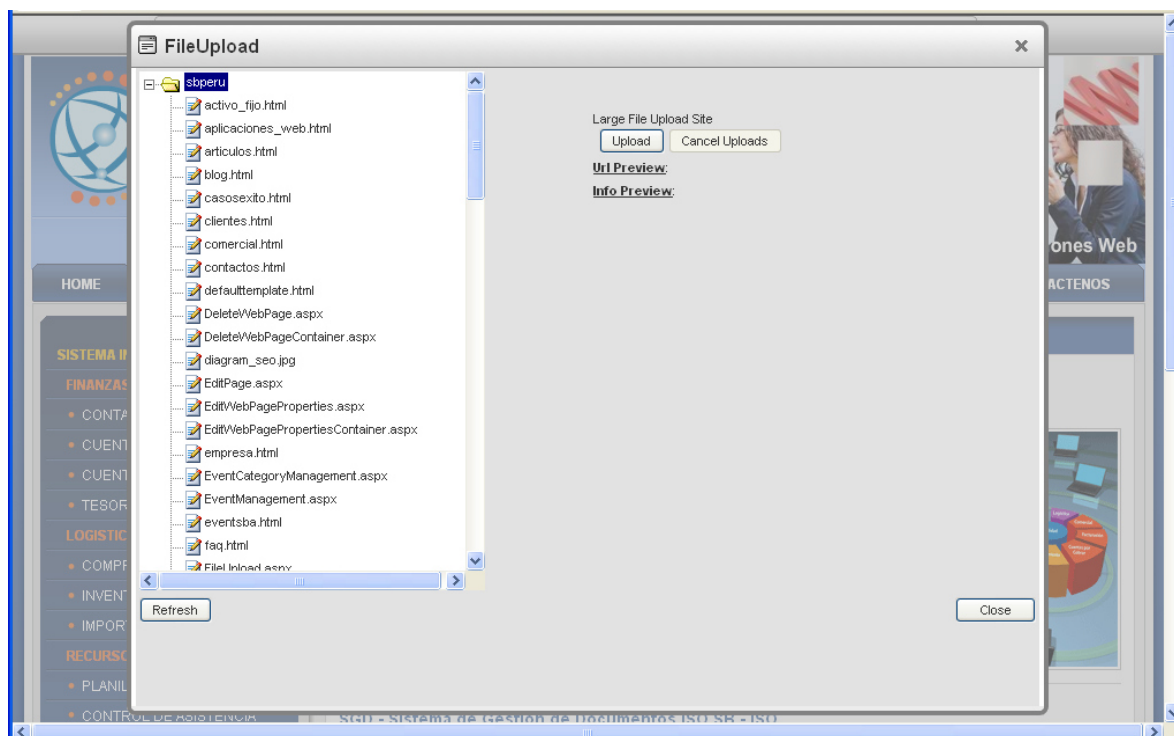


Figura 6.20: Cargar un archivo en el Portal Web.

b. Administración del Menú Principal

Permite agregar, modificar y eliminar opciones del Menú del Portal Web, a fin de mejorar la gestión de la navegabilidad del portal Web que favorece la optimización SEO, ver la figura 6.21. Esta opción permite:

- Agregar Nuevo / Modificar
- Eliminar

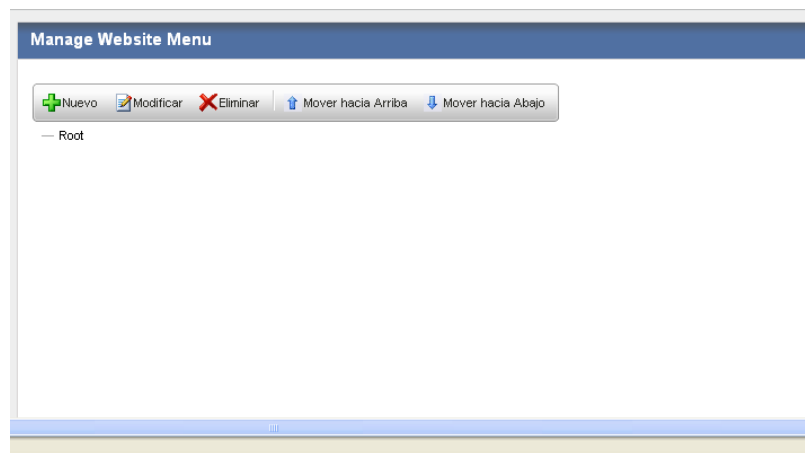


Figura 6.21: Gestión del Menú Principal.

c. Mapa de Sitio

Permite agregar, modificar y eliminar elementos del Mapa de Sitio que darían mayor facilidad para la gestión de Navegabilidad del Portal Web, ver la figura 6.22. Esta opción permite:

- Nuevo / Modificar
- Eliminar



Figura 6.22: Gestión del Mapa de Sitio.

d. Robots

Crea y modifica el archivo robots.txt. El usuario administrador puede crear y modificar el archivo robot.txt a fin de controlar la navegabilidad del portal.

- **Opción: Ayuda SEO**

Permite disponer de una guía acerca de cómo usar las mejores técnicas SEO para mejorar el posicionamiento en los Motores de Búsqueda.

En la figura 6.23 se muestra como la Ayuda SEO está disponible como una opción del Menú.



Figura 6.23: Ayuda SEO.

6.3 Elección del Caso de Estudio

Se ha escogido como Caso de Prueba, el portal Web de Soft Brilliance S. A. C., el cual es una empresa peruana que ofrece servicios de desarrollo de sistemas empresariales.

La elección de este portal se debe a que Soft Brilliance S. A. C., al igual que otras pequeñas y medianas empresas peruanas, no están adecuadamente posicionadas en los principales Motores de Búsqueda, a pesar de contar con páginas de diseños atractivos.

Soft Brilliance S. A. C., viene dando sus servicios desde hace 6 años y hace 3 decidió implementar su propio portal Web a manera de publicidad en la Internet. En ella darían a conocer sus principales servicios, así como sus principales clientes.

Pero una encuesta realizada por la empresa demostró que la mayor parte de clientes llegan a saber de los servicios mediante un directorio impreso de empresas, en la cual está publicada la URL del Portal. En conclusión, una mínima parte de los nuevos clientes llegan a conocer la empresa a través de los buscadores Web.

El que los buscadores no muestran el portal, se debe a que probablemente no fue diseñado pensando en el posicionamiento Web y menos en las recomendaciones SEO.

6.3.1 Análisis del Estado Inicial del Caso de Estudio

Una vez elegido el caso de estudio, es necesario conocer el estado actual del portal, identificando algunos problemas o deficiencias, que evitan sea posicionado en los motores de búsqueda.

Las pruebas realizadas fueron hechas en base a las recomendaciones SEO. A continuación, se enumeran los objetivos de esta prueba:

- Comprobar si el portal está en los motores de búsqueda
 - Examinar las páginas para encontrar problemas
 - Determinar en qué posiciones es mostrado el portal Web
-
- **Comprobar si el portal Web es indexado por los principales Motores de búsqueda**

Lo primero que debemos verificar es si el portal está realmente en un motor de búsqueda o directorio. Es decir, si aparece cuando alguien lo busca en Google o en el directorio de Yahoo.

Para esta prueba, se usará Google ya que es el buscador más importante.

a. Examinar en la caché de los buscadores

Para acceder a la caché de los buscadores, primero se buscó en el buscador Google la página principal de SoftBrilliance, luego se seleccionó la opción ‘En Caché’, tal como se muestra en la figura 6.24.

Finalmente, el buscador muestra el registro en su caché de la página del portal. En la figura 6.25 se ve que efectivamente el buscador conoce la página principal de Softbrilliance.

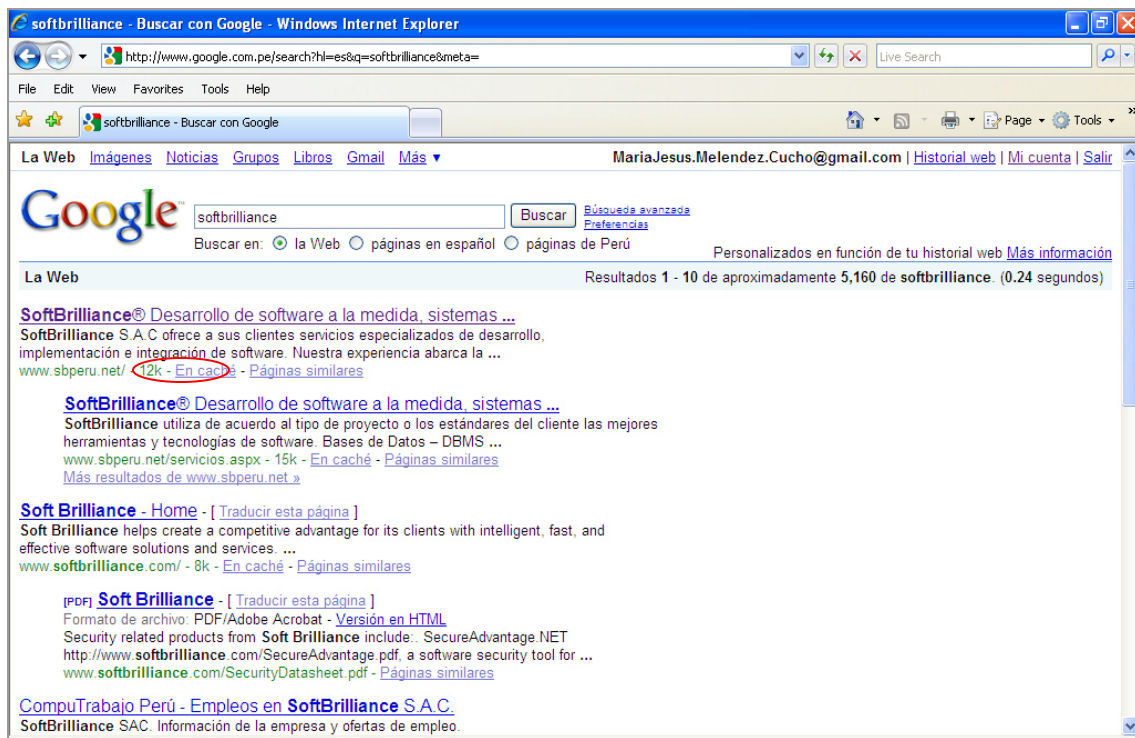


Figura 6.24: Buscar página principal del portal

**SOFTBRILLIANCE**
SOFTWARE DEVELOPMENT

HOME EMPRESA CLIENTES SOLUCIONES SERVICIOS CASOS DE ÉXITO BLOGS CONTACTÉNOS

Menú de Sistemas

- ▶ SISTEMA INTEGRADO SBA
 - ▶ Contabilidad
 - ▶ Inventarios
 - ▶ Compras
 - ▶ Ventas
 - ▶ Tesorería
 - ▶ Cuentas por Pagar
 - ▶ Activos Fijos
 - ▶ Producción
 - ▶ Mantenimiento
 - ▶ Planillas
 - ▶ Control de Asistencia
- ▶ SISTEMA GESTIÓN DE RECLAMOS
- ▶ SISTEMA DE GESTIÓN DOCUMENTOS ISO
- ▶ SISTEMA DE GESTIÓN ARCHIVOS DOCUMENTARIOS

Menú Adicional

Bienvenido a www.sbperu.net gracias por su visita.

Fecha | Lunes, 16 de Marzo del 2009

**SBA**
Sistema Integrado

Sistema de Gestión de documentos ISO

Sistema de Gestión de Archivos Documentarios

SOLUCIÓN TOTAL AL PROBLEMA DE SUS PLANILLAS
SMART BUSINESS APPLICATION – SBA V 4.0 MODULO RR.HH.
INCORPORA: RTPS y CONTROL DE ASISTENCIA

SoftBrilliance S.A.C. ofrece a sus clientes servicios especializados de desarrollo, implementación e integración de software. Nuestra experiencia abarca la automatización de procesos en: Finanzas, Logística, Comercialización, Manufactura, Contabilidad, Recursos Humanos, etc. Tenemos la capacidad demostrada de entender la problemática particular de su empresa y brindarle el servicio personalizado y eficiente que está buscando.

Nuestro principal respaldo es la cartera de clientes conformada por empresas peruanas y extranjeras de primer nivel entre las que destacan: *Market Flex, Banco de Crédito, Neptunia, Packaging Products, TGP, BACKUS, PECSA GAS*. Todas ellas han confiado plenamente en nosotros para hacer realidad los proyectos de sistemas más desafiantes.

➤ Información respecto a nuestros productos y servicios, escriba a info@sbperu.net

COPYRIGHT © 2001 - 2007 **SOFTBRILLIANCE** S.A.C. | ALL RIGHT RESERVED | RELEASE 3.0

Figura 6.25: Como se muestra en la cache de Google el Portal Soft Brilliance S.A.C.

b. Ver el número de páginas indexadas

Luego de verificar si el portal es indexado, se desea saber cuantas de sus páginas son indexadas. Para lo cual se ingresa el texto ‘-pppp site: www.sbperu.net’ en el buscador. La prueba mostró que son 56 páginas de las cuales tiene conocimiento el buscador, tal como lo muestra la figura 6.26.

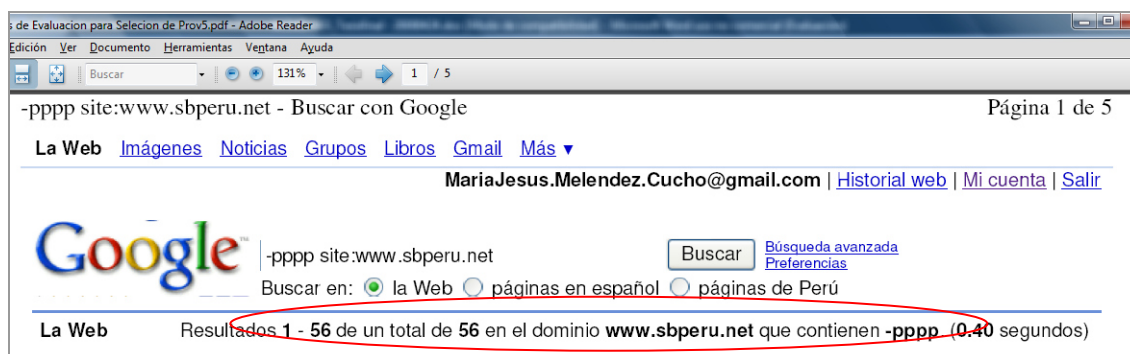


Figura 6.26: Como se muestra en la cache de Google el Portal Soft Brilliance S.A.C.

En resumen, de acuerdo a las dos pequeñas pruebas realizadas, podemos concluir que Google indexa y tiene conocimiento de las páginas del portal de Softbrilliance.

Se cumplió el primer objetivo, el de comprobar que Google conoce el portal de SoftBrilliance. Pero esto no es prueba suficiente de un buen posicionamiento, por lo cual, se analizará las páginas Web siguiendo las más importantes recomendaciones SEO.

• Comprobar la optimización de las Páginas Web

Este análisis se hizo sobre el código html de las páginas del portal Web a fin de encontrar algunas deficiencias que impide tener páginas optimizadas.

Luego del análisis respectivo, se encontraron faltas a las normas para la optimización de páginas Web, entre las cuales destacan:

a. La posición del TAG TITTLE es incorrecto.

En el código fuente de la página principal, el TAG TITTLE no se encuentra inmediatamente después del TAG HEAD, tal como se muestra en la figura 6.27.

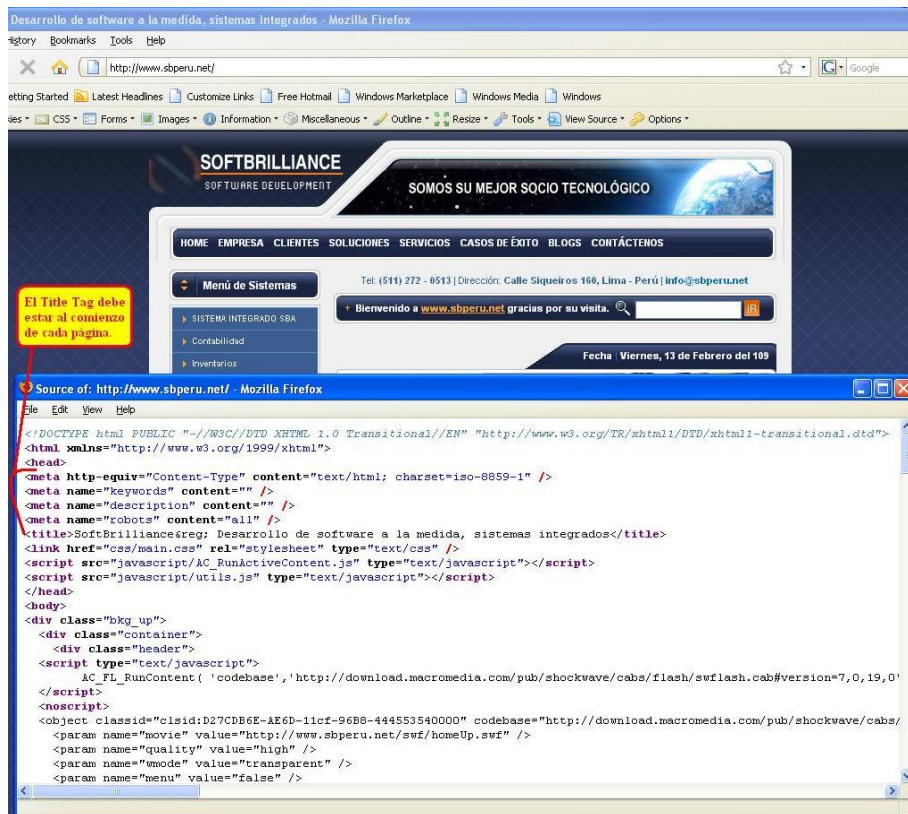


Figura 6.27: La mala ubicación del TAG TITTLE.

b. La información en el TAG TITTLE se repite en todas las páginas del portal.

Todas las páginas tienen el mismo texto en los TAG TITTLE. Pero SEO considera que una página Web debe tener un TITTLE diferente y significativo, con alrededor de unos 70 caracteres, en el cual deben encontrarse las palabras clave.

En la figura 6.28 se muestran 2 páginas del portal con el mismo TITTLE.

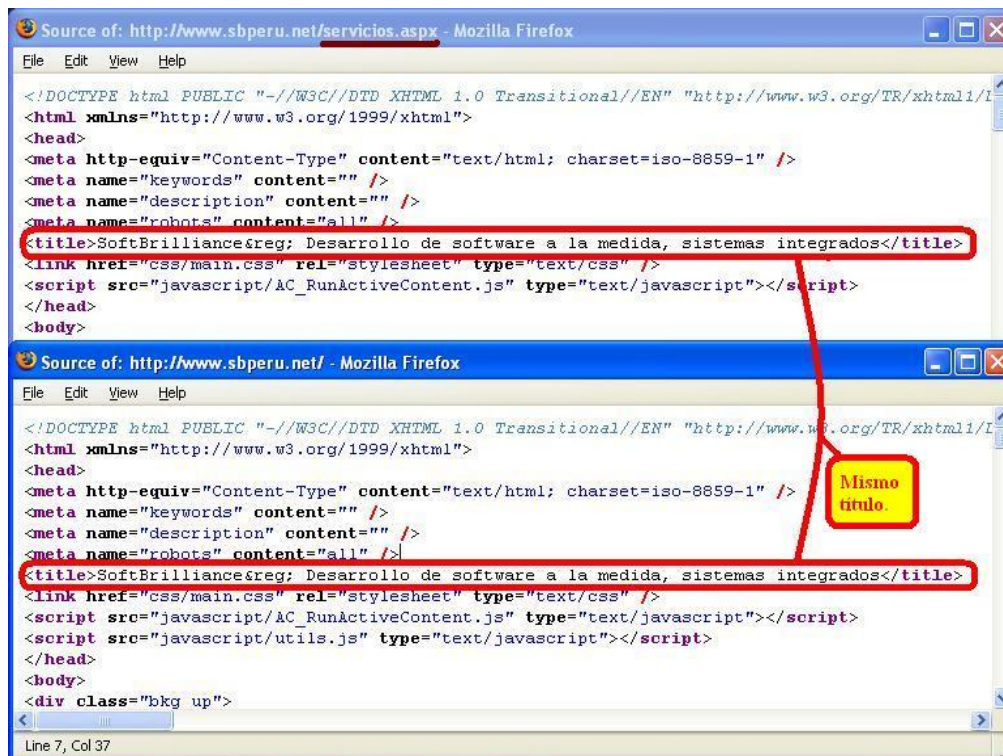


Figura 6.28: El texto en el TAG TITTLE se repite en la página de Inicio y en la página 'servicios.aspx'.

c. Las páginas Web no tiene el METATAG DESCRIPTION, ni el METATAG KEYWORD

En la figura 6.28 también se puede ver que las páginas no cuentan con los METATAG de mayor relevancia como son los DESCRIPTION Y LOS KEYWORD.

d. Varias partes del Sitio Web fueron hechas en Flash

Muchas de las páginas del portal poseen grandes imágenes flash, pero la existencia de estos plantea problemas de cara a los motores de búsqueda, pues difícilmente podrán acceder al contenido, aunque se sabe que Google empieza a leer el contenido de los archivos flash, no se conoce aún la forma de tratarlos, en cualquier caso, no es acertado pues plantea problemas de accesibilidad.

En las figuras 6.29 y 6.30 se muestran los principales archivos flash que posee el portal de Softbrilliance.



Figura 6.29: Archivo Flash que se muestran en la página principal del portal



Figura 6.30: Archivo Flash que se muestran en la cabecera de las páginas del portal

e. No se tienen palabras claves en los encabezados H1, H2

Respecto a encabezados H1, H2 o similares no existen, palabras clave que puedan ser de interés, tal como se muestra en las figuras 6.31 y 6.32.

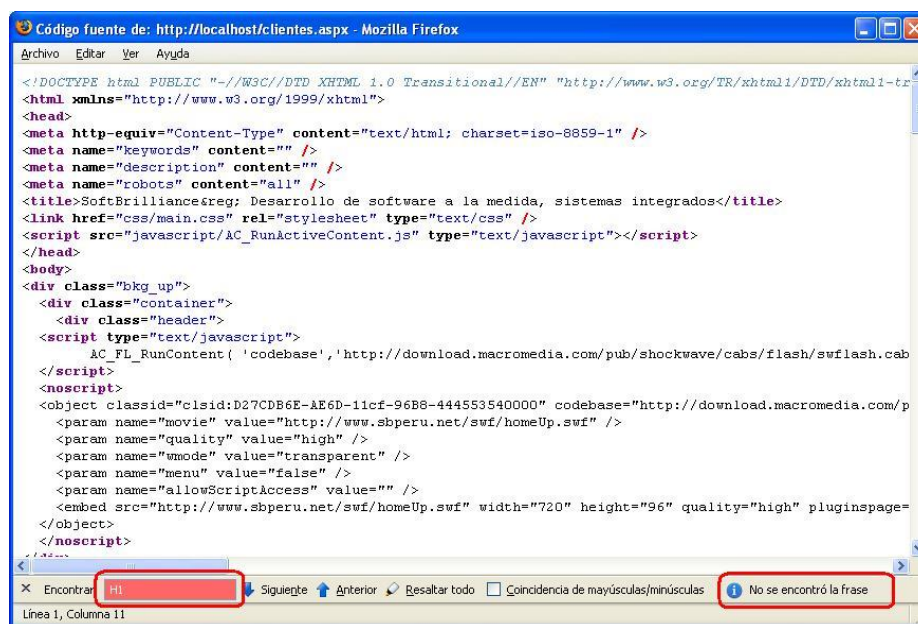


Figura 6.31: La página cliente.aspx no cuenta con texto en el encabezado H1.

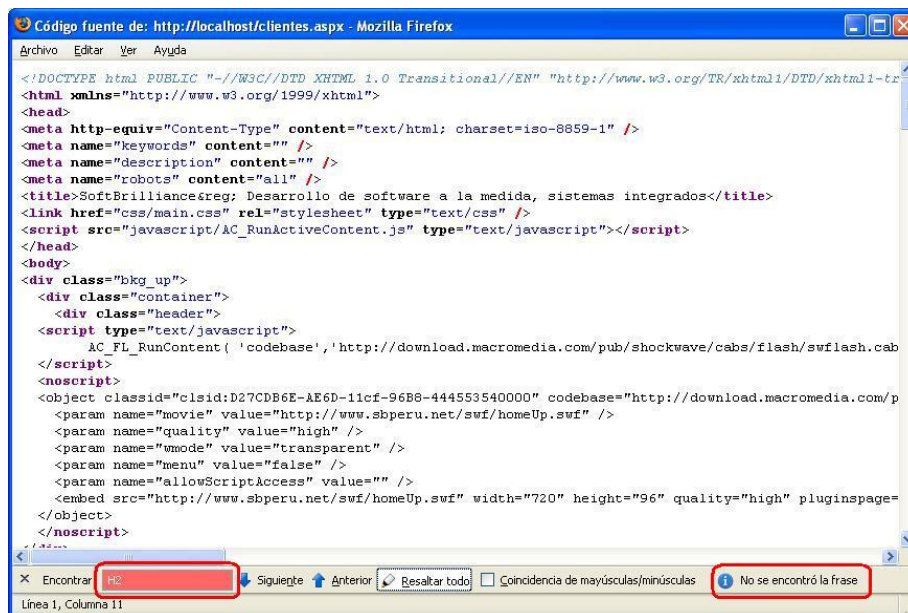


Figura 6.32: La página cliente.aspx no cuenta con texto en el encabezado H1.

f. Los atributos ALT no están optimizados

Los atributos ALT de las imágenes no están optimizados, es decir no brindan ningún tipo de información acerca de la imagen. Es mejor asociar un texto informativo a cada imagen.

En la figura 6.33 se muestra las imágenes del portal que no tienen el atributo ALT.



Figura 6.33: La imagen de la página no tiene el atributo ALT.

En resumen, de acuerdo a las pruebas realizadas en los códigos fuentes, podemos concluir que el portal Web no está teniendo en cuenta criterios básicos (recomendaciones SEO) para la optimización de las páginas Web.

- **Comprobar propiedades del portal Web**

Otra de las pruebas es comprobar si el portal posee o maneja algunos elementos importantes para los motores de búsqueda.

- a. El portal Web no dispone del archivo robots.txt**

El portal Web no cuenta con el archivo robots.txt, el cual tiene importancia en la navegabilidad de las páginas por los motores de búsqueda.

En la figura 6.34 se puede ver que el portal no contiene dicho archivo.

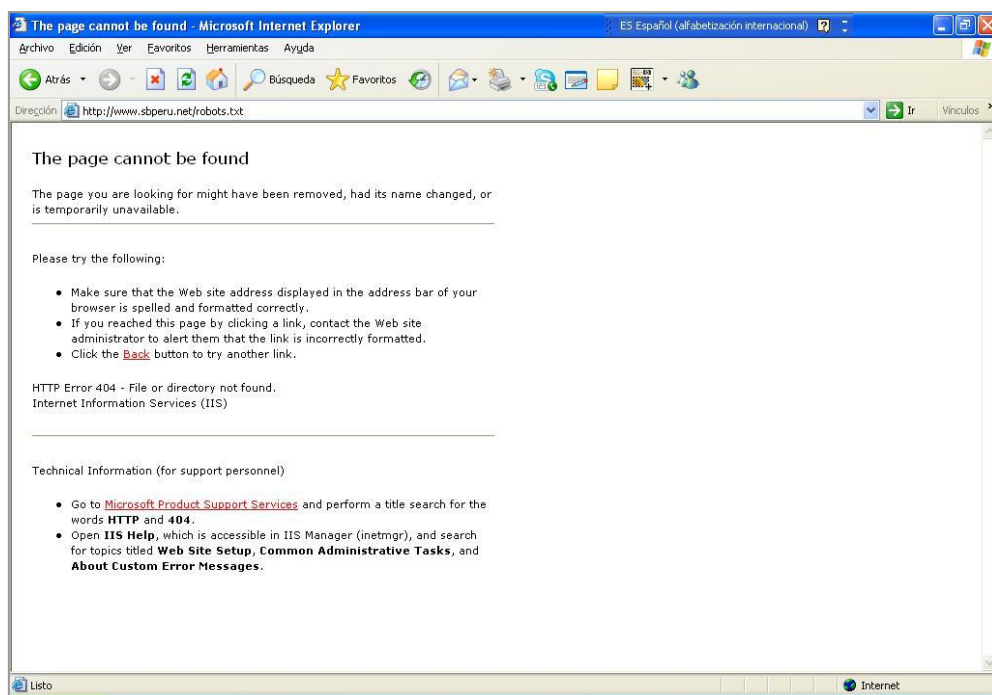


Figura 6.34: El portal Web no cuenta con el archivo robots.txt.

- b. El portal Web no tienen un sitemap.xml**

El portal Web no cuenta con el archivo sitemap.xml, el cual tiene importancia en la navegabilidad de las páginas por los motores de búsqueda.

En la figura 6.35 se puede ver que el portal no contiene dicho archivo.

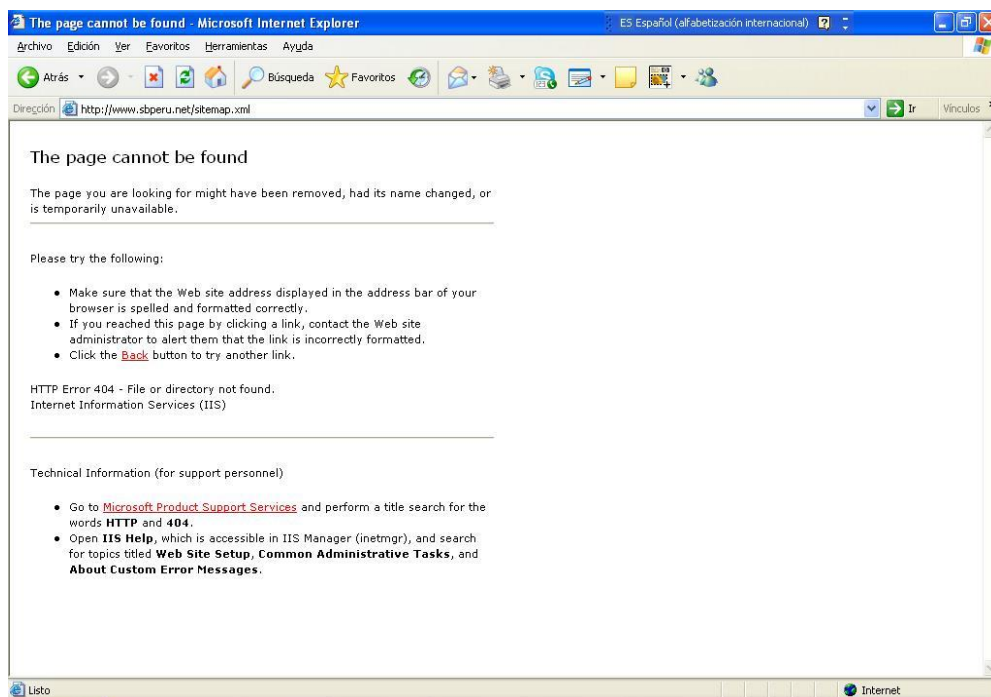


Figura 6.35: El portal Web no cuenta con el archivo sitemap.xml.

c. El portal Web no cuenta con una canonización de URLs

Uno de los problemas del portal es la no canonización de URLs. Es decir se muestra la misma página para las siguientes direcciones: www.sbperu.net, www.sbperu.net/index.aspx, sbperu.net y sbperu.net/index.aspx. El problema radica en que los Motores de Búsqueda registran las direcciones como si fueran distintas.

Luego de mostrar las faltas en las propiedades del portal, se puede resumir que el portal Web no cuenta con los archivos necesarios para la navegabilidad como son: el robots.txt y el sitemap.xml.

Los resultados de los diferentes análisis hechos al portal de Softbrilliance está dando más idea del estado inicial de las páginas Web, que justificarían el bajo posicionamiento en los buscadores.

Ahora, veamos como los Motores de Búsqueda están indexando el portal Web de SoftBrilliance. A continuación, se muestra los resultados de los análisis, que se realizaron en base a reportes de medición de las posiciones del portal Web con relación a diferentes Palabras Claves.

- **Comprobar el posicionamiento inicial en la que es mostrado el portal Web de SoftBrilliance S. A. C.**

Para poder visualizar mejor el posicionamiento del portal de SoftBrilliance, se generó reportes en base a determinadas palabras claves, las cuales son con la que la mayoría de usuarios visitantes intentan ubicar al portal.

Se escogió alrededor de 25 palabras claves, mediante las cuales se mide el posicionamiento para dicha palabra. También se escogieron para la prueba, los 3 motores de búsqueda más utilizados, como son: Google, Yahoo y MSN.

A continuación, se muestra el posicionamiento en cada buscador mencionado. El análisis más completo, que incluye los respectivos reportes, se adjuntó en el Anexo B.

a. Posicionamiento en Google

Los reportes mostraron que Google llega a posicionar al portal de SoftBrilliance en la primera página, pero con sólo un 12 % de palabras claves, tal como se muestra en la figura 6.36. Por otro lado, en la figura 6.37 se muestra la proporción de palabras claves, mediante las cuales, el portal es posicionado.

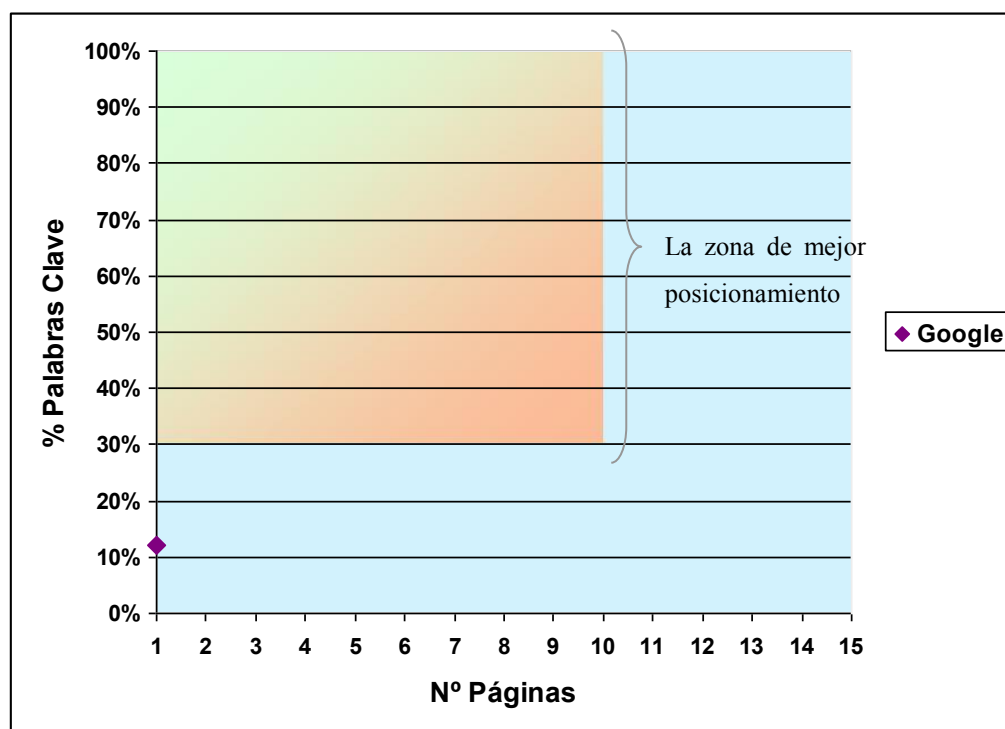


Figura 6.36: Puntos de posicionamiento del portal Web en Google.

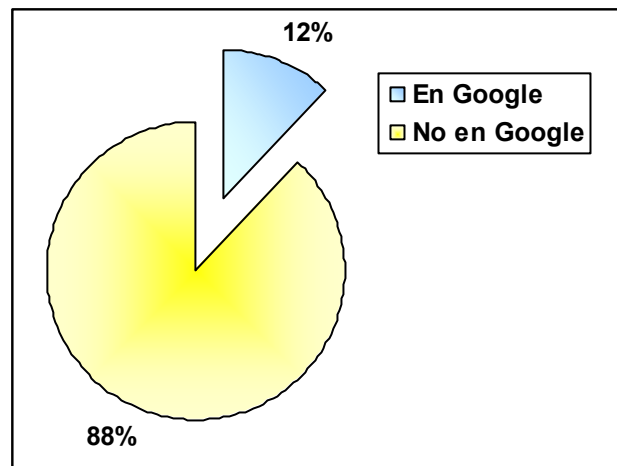


Figura 6.37: El porcentaje de palabras claves por las que es posicionado en Google.

En la figura 6.36, se puede observar la zona de mejor posicionamiento, esta región muestra que un adecuado posicionamiento es estar al menos con el 30 % de palabras claves en las primeras 10 páginas.

Teniendo definido la zona de un adecuado posicionamiento, se puede concluir, que el posicionamiento en Google no es el adecuado, ya que solo es posicionado con el 12 % de palabras claves, los cuales además son palabras con el nombre de la empresa más no expresan el rubro o la actividad comercial.

b. Posicionamiento en Yahoo

Al igual que con Google, se generó reportes a fin de mostrar como Yahoo llega a posicionar al portal de SoftBrilliance. El resultado mostró que Yahoo posiciona al portal en la primera página con un 12 % de palabras claves pero además lo ubica en la segunda página con el 4% de palabras claves, tal como se muestra en la figura 6.38.

Por otro lado, en la figura 6.39 se puede observar la proporción de palabras claves, por el cual, el portal es posicionado.

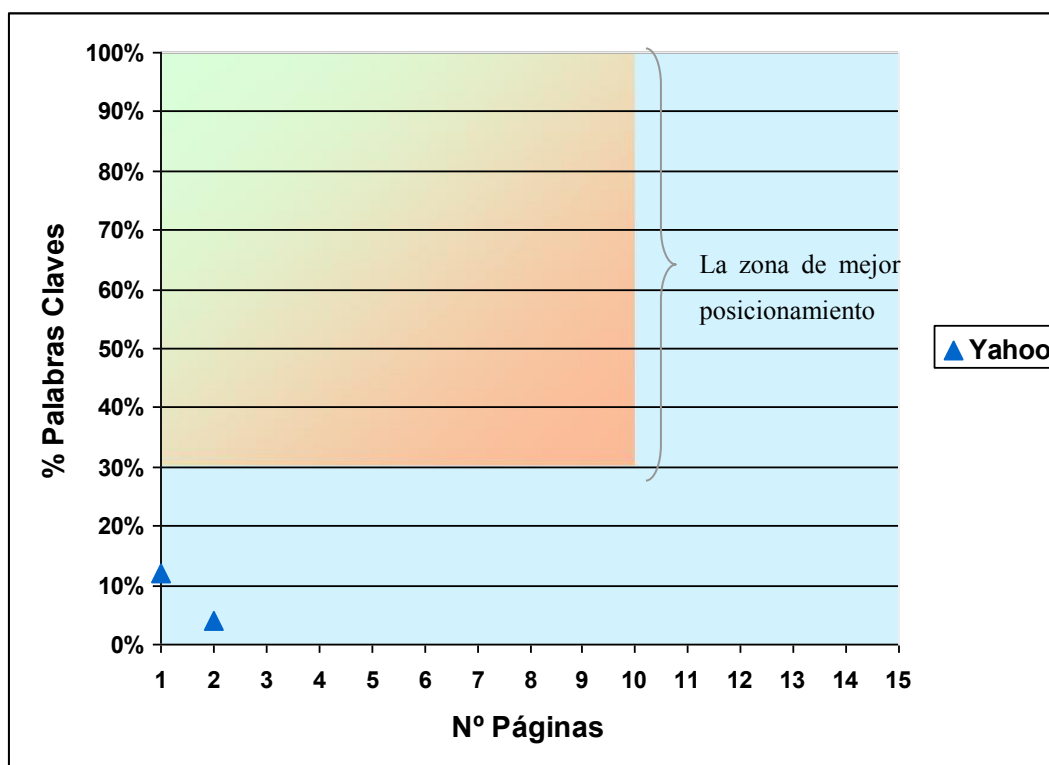


Figura 6.38: Punto de posicionamiento del portal Web en Yahoo.

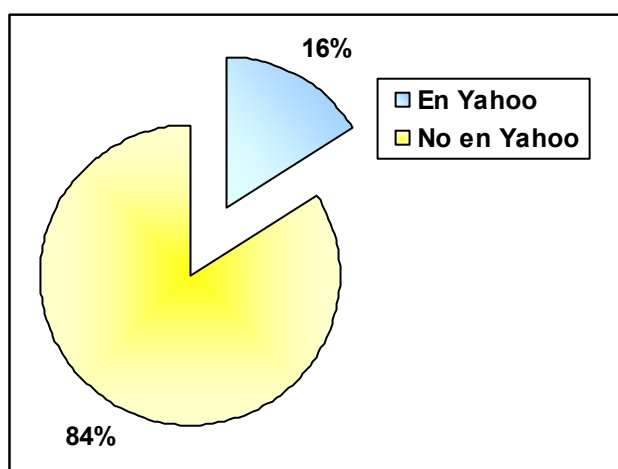


Figura 6.39: El porcentaje de palabras claves por las que es posicionado en Yahoo.

Teniendo definido la zona de un adecuado posicionamiento, se puede concluir, que el posicionamiento en Yahoo tampoco es el adecuado, ya que solo con un 16 % de palabras claves, es posicionado en las dos primeras páginas.

Si bien es cierto está en las dos primeras páginas, se debe especificar que, las palabras claves por la cual es posicionado, tienen en su mayoría, el nombre de la empresa más no palabras que describen el rubro o la actividad de la empresa.

c. Posicionamiento en MSN

Al igual que los otros Motores de búsqueda, se generó reportes de cómo MSN llega a posicionar al portal de SoftBrilliance, como se muestra en la figura 6.40. El resultado mostró que:

- A partir de la posición 5 de la primera página con un 12 % de palabras claves.
- En la cuarta página con el 8% de palabras claves.
- En la séptima página con el 4% de palabras claves.

Además, en la figura 6.41 se puede observar la proporción de palabras claves, por el cual, el portal es posicionado.

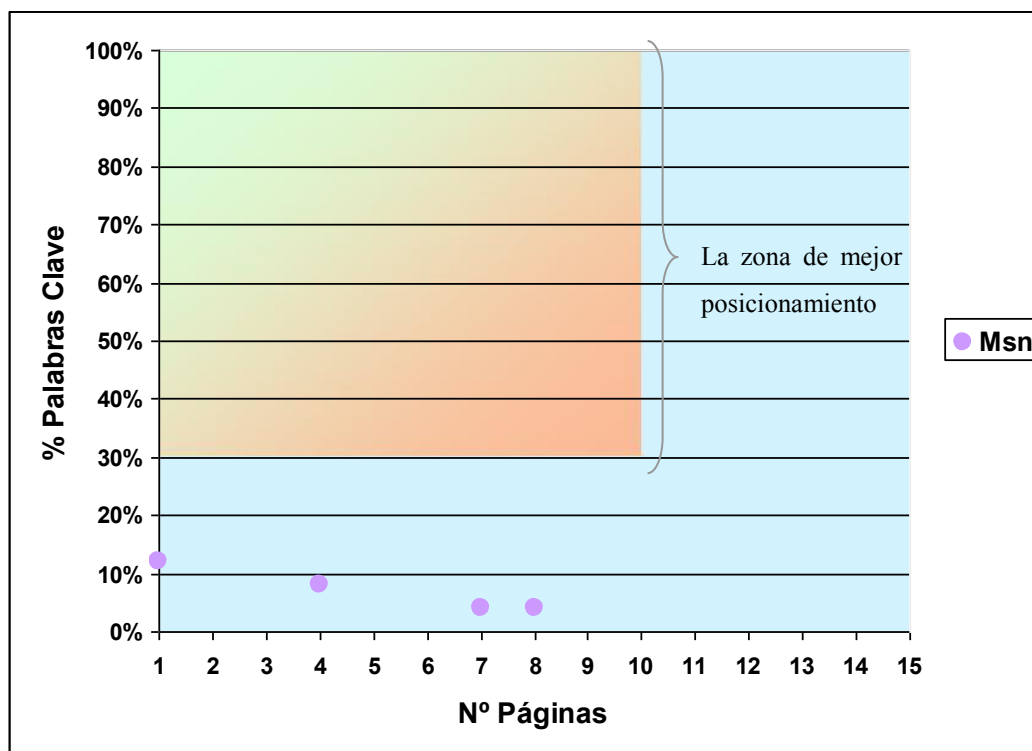


Figura 6.40: Punto de posicionamiento del portal Web en MSN.

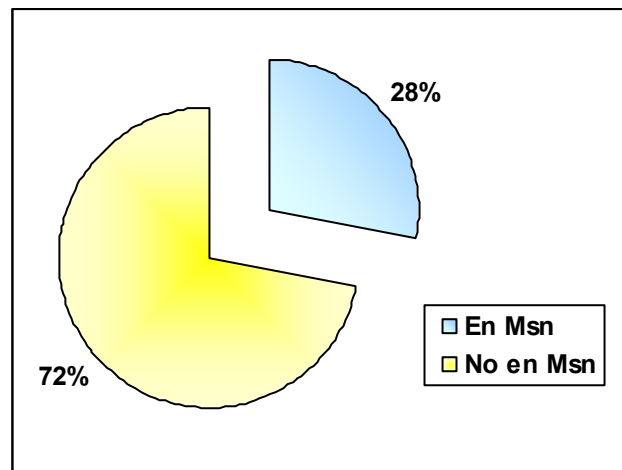


Figura 6.41: El porcentaje de palabras claves por las que es posicionado en MSN.

Teniendo definido la zona de un adecuado posicionamiento, se puede concluir, que el posicionamiento en MSN tampoco es el adecuado, ya que con un 28 % de palabras claves, es posicionado en las ocho primeras páginas.

Si bien es cierto, MSN posiciona el portal en varias páginas pero las palabras claves por la cual es posicionado, tienen en su mayoría, el nombre de la empresa más no palabras que describen el rubro o la actividad de la empresa.

d. Posicionamiento en los tres Motores de Búsqueda

A manera de saber el posicionamiento promedio en estos tres proveedores, se generó unos resultados conjuntos:

- A partir de la posición 5 de la primera página con un 12 % de palabras claves.
- En la segunda página con el 0.33% de palabras claves.
- En la cuarta página con el 2.67% de palabras claves.
- En la séptima página con el 1.33% de palabras claves.
- En la octava página con el 1.33% de palabras claves.

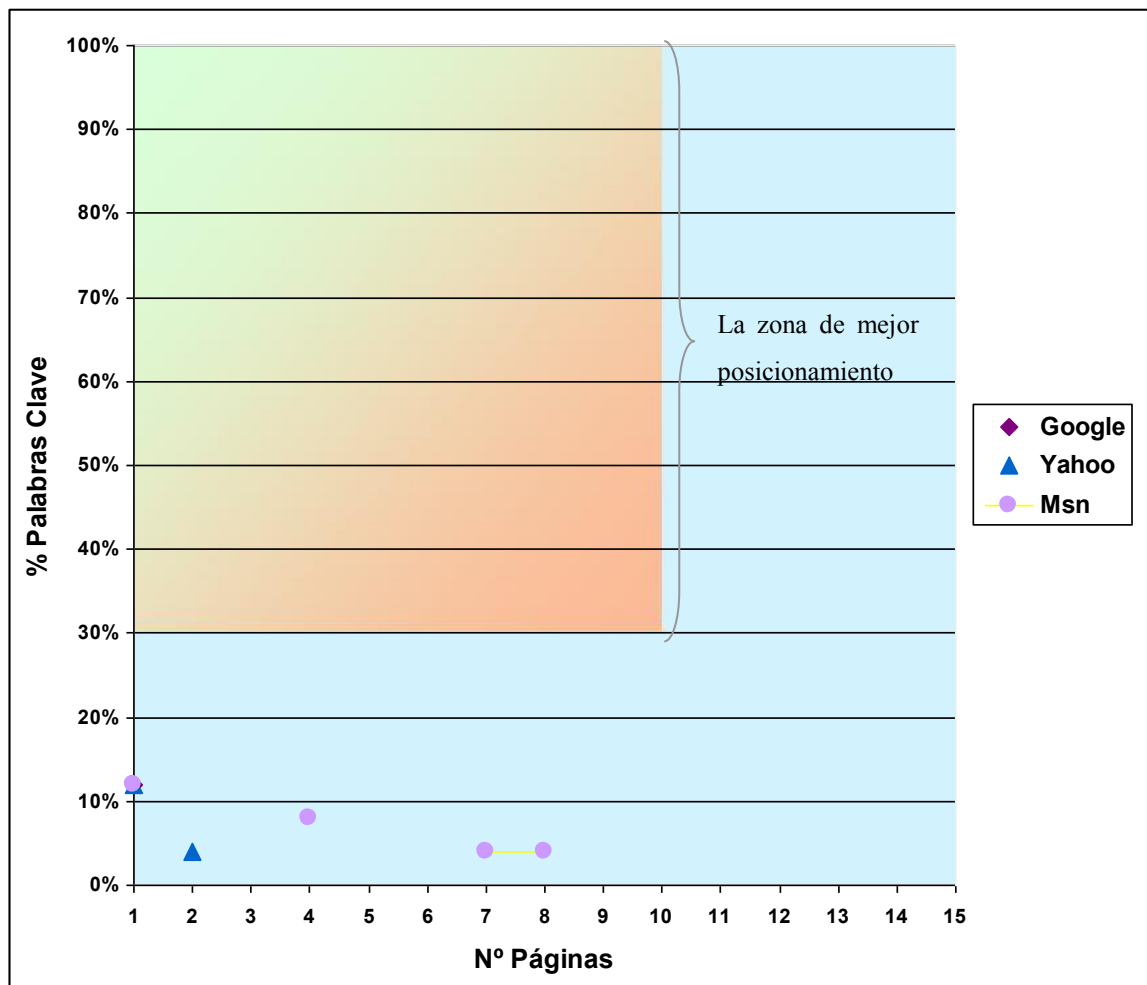


Figura 6.42: Puntos de posicionamiento de los tres principales Motores de Búsqueda.

Los posicionamientos en los diferentes Motores de Búsqueda, se localizan por debajo del 20% de palabras claves, es decir, sólo una pequeña parte de estas palabras logra ser indexado, vea la figura 6.42.

En los tres Motores de Búsqueda son posicionados en algunas páginas, pero con aquellas palabras claves que representan el nombre de la empresa, pero se necesita que esté posicionado con más palabras, que representen los rubros o actividades de la empresa.

- **Resumen del Análisis del estado inicial**

Por todo lo visto anteriormente, el análisis de las páginas y los reportes de los motores de búsqueda, se puede llegar a las siguientes conclusiones:

- Se desconoce los métodos de posicionamiento en buscadores, ya que el portal Web esta siendo indexado por los motores de búsqueda pero de forma automática. Por otro lado el portal solo es conocido por la marca (nombre de la empresa), con lo cual podemos decir que no seria fácil ubicarlo en Internet.
- No existe una estrategia de marketing hacia el portal Web, simplemente se decidió construirlo con la idea de poseer uno de ellos en la Internet, pero no se ha visto como puede llegar a ser mas útil o como conseguir más usuarios visitantes a través de ella.
- El portal Web es amigable para el usuario visitante (atractivo), pero hay ciertas deficiencias en cuanto a ser amigables a los motores de búsqueda. Ya que su posicionamiento en los principales motores de búsqueda es mínima.

6.3.2 Aplicando la Solución Propuesta al Caso de Estudio

Luego de haber analizado el estado inicial del Portal Web, se decidió implantar el prototipo propuesto en dicho Portal, a fin de demostrar que la aplicación de la solución propuesta permitirá mejorar el posicionamiento en los principales motores de búsqueda. A continuación se describen, paso a paso, los cambios efectuados en el Portal de SoftBrilliance S. A. C:

- **Nuevo Diseño del Portal Web**

En primer lugar se aplicó un nuevo diseño para el Portal Web de SoftBrilliance S. A. C. (Página principal y sus páginas internas). Este nuevo diseño se basaba en una nueva estructura html, el cual resulta ser más amigable para los motores de búsqueda. En la figura 6.43 y 6.44 se muestran el nuevo diseño del portal.



Figura 6.43: Nuevo Diseño de la Página Principal del Portal Web de SoftBrilliance S. A. C.

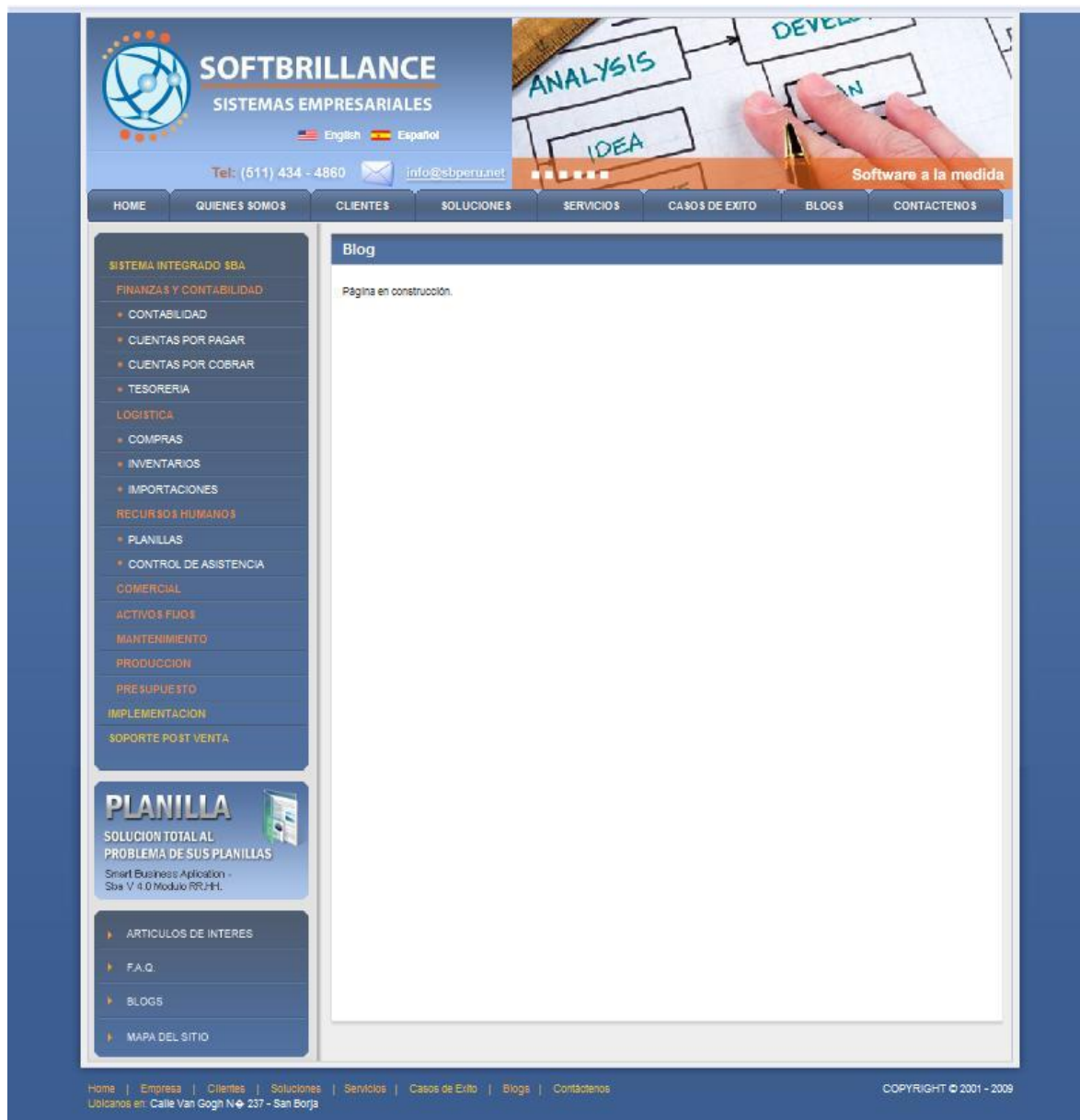


Figura 6.44: Nuevo Diseño de las Páginas Internas del Portal Web de SoftBrilliance S. A. C.

Se diseñó una “plantilla base”, para la creación de nuevas páginas Web dentro del portal, esta plantilla fue construida considerando las formas de optimización SEO (véase la figura 6.45).

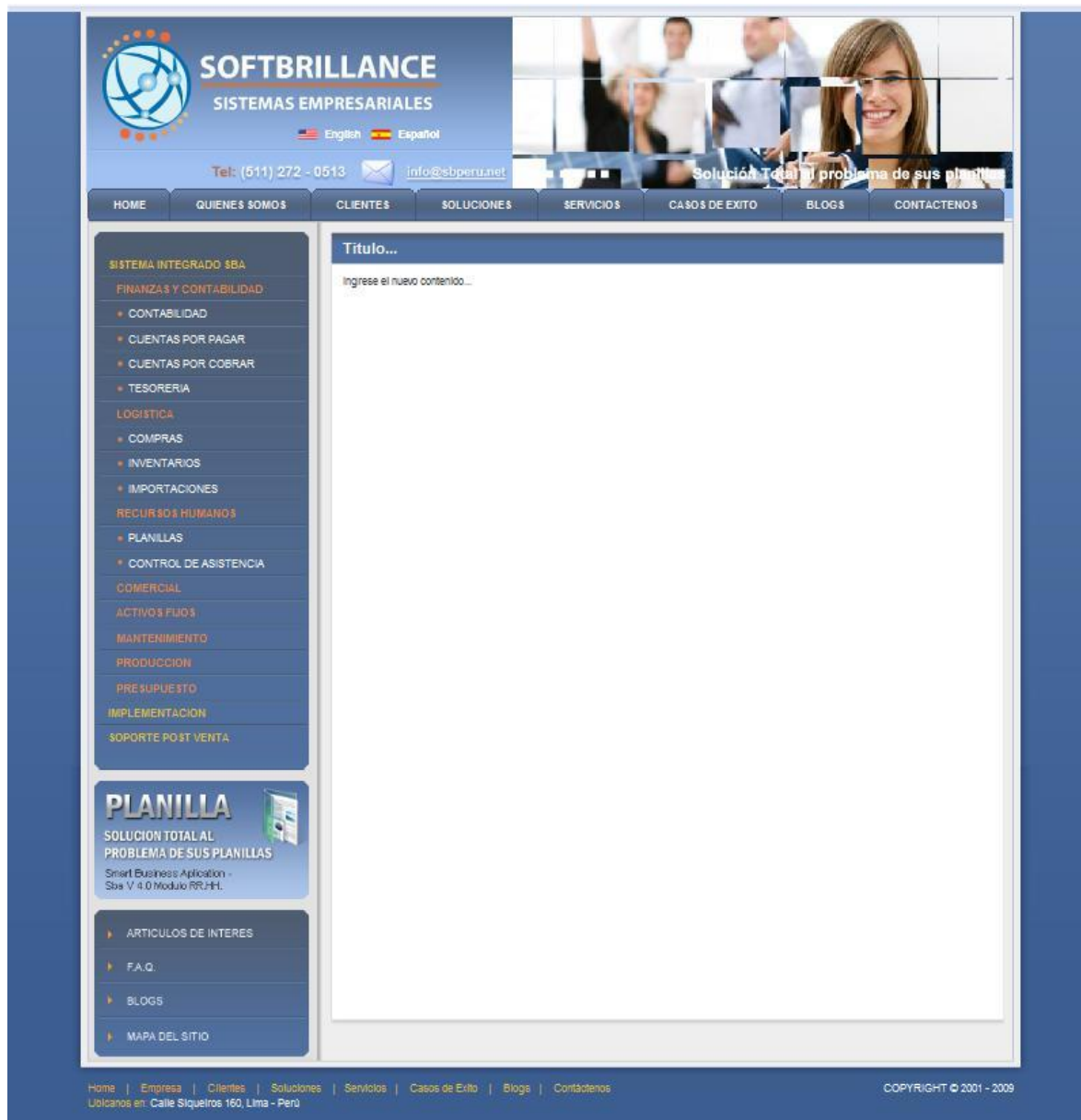


Figura 6.45: Plantilla Base orientado a SEO

- **Configuración de una única URL para el portal**

Se necesita que el portal sea reconocido por una única Url de dominio tanto por los usuarios navegantes como por los motores de búsqueda. Para esto se implementó la canonización de Urls, el cual es www.sbperu.net

A continuación, se listan los pasos realizados para la canonización.

Primera Canonización: Busca que el portal Web sea reconocido por varias páginas de inicio estándar, los cuales son conocidos por la mayoría de usuarios navegantes, como por ejemplo: inicio.html, default.aspx, index.htm, principal.asp, etc.

A continuación se muestra la tabla 6.1, con un listado de posibles páginas de inicio de los portales Web.

Paginas de Inicio Estándar	Extensión
index	.htm
	.html
	.shtm
	.shtml
	.asp
	.aspx
default	.htm
	.html
	.shtm
	.shtml
	.asp
	.aspx
inicio	.htm
	.html
	.shtm

	.shtml
	.asp
	.aspx
home	.htm
	.html
	.shtm
	.shtml
	.asp
	.aspx
main	.htm
	.html
	.shtm
	.shtml
	.asp
	.aspx
principal	.htm
	.html
	.shtm
	.shtml
	.asp
	.aspx

Tabla 6.1: Listado de Páginas de Inicio Estándar en los Portales Web.

Segunda Canonización: Busca el uso del ‘www’ al inicio del dominio, a pesar de que los usuarios navegantes no ingresan este. A continuación algunos ejemplos:

- <http://sbperu.net/> → <http://www.sbperu.net/>
- <http://sbperu.net/clientes.html> → <http://www.sbperu.net/clientes.html>
- <http://sbperu.net/index.html> → <http://www.sbperu.net/>

Tercera Canonización: Esta tercera forma de canonización es para cuando se desea ingresar por el IP Address del Portal Web.

- <http://64.79.161.44/> → <http://www.sbperu.net/>

- **Redirección de páginas antiguas para no perder el tráfico de los buscadores**

Se implementó el redirect 301, para aquellas páginas del Portal, que ya no se desea que se visualicen su contenido para los usuarios visitantes en los motores de búsqueda, e intercambiarlo por un contenido actual. A continuación se muestra algunos ejemplos de redirecciones 301:

- <http://www.sbperu.net/soluciones.aspx> → <http://www.sbperu.net/soluciones.html>
- <http://www.sbperu.net/servicios.aspx> → <http://www.sbperu.net/servicios.html>
- <http://www.sbperu.net/casosexito.aspx> → <http://www.sbperu.net/casosexito.html>

Este redirect 301, está basado en un archivo xml, el cual contiene las Urls “de” (*from*) y las Urls “a” (*to*). A continuación se muestra, en la figura 6.46, el archivo XML del Portal Web de SoftBrilliance S. A. C.

Este fichero XML no parece tener ninguna información de estilo asociada. Se muestra debajo el árbol del documento.

[illegible]

Figura 6.46: Archivo XML para el redirect 301.

- **Modificación de Propiedades de la Página Web**

Se ingresó valores en cada una de las propiedades de las páginas, tales como los TITLE, METATAGs, INDEX, FOLLOW, etc...

Además, Se ingresó contenidos, imágenes, todos basados en palabras claves para la optimización del portal, véase la figura 6.47.

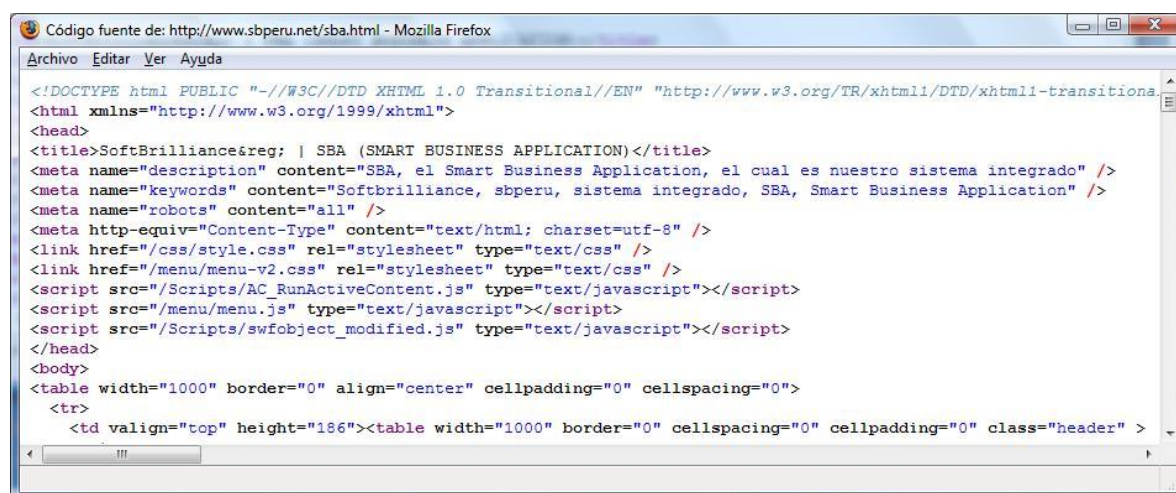


Figura 6.47: Ingreso de palabras claves en las propiedades principales de la página.

- **Archivo robots.txt**

Se creó el archivo robots.txt, con el cual hicimos que ciertas páginas o fólderes del Portal Web, no sean accesibles por los robots de los motores de búsqueda, con la finalidad de que éstos no puedan indexar contenido no útil para usuarios navegantes. Véase la figura 6.48.

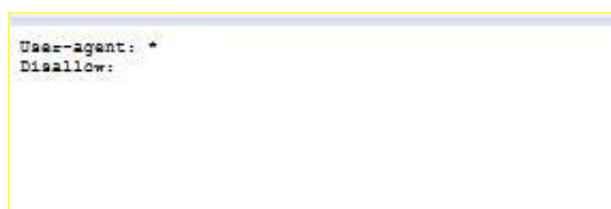
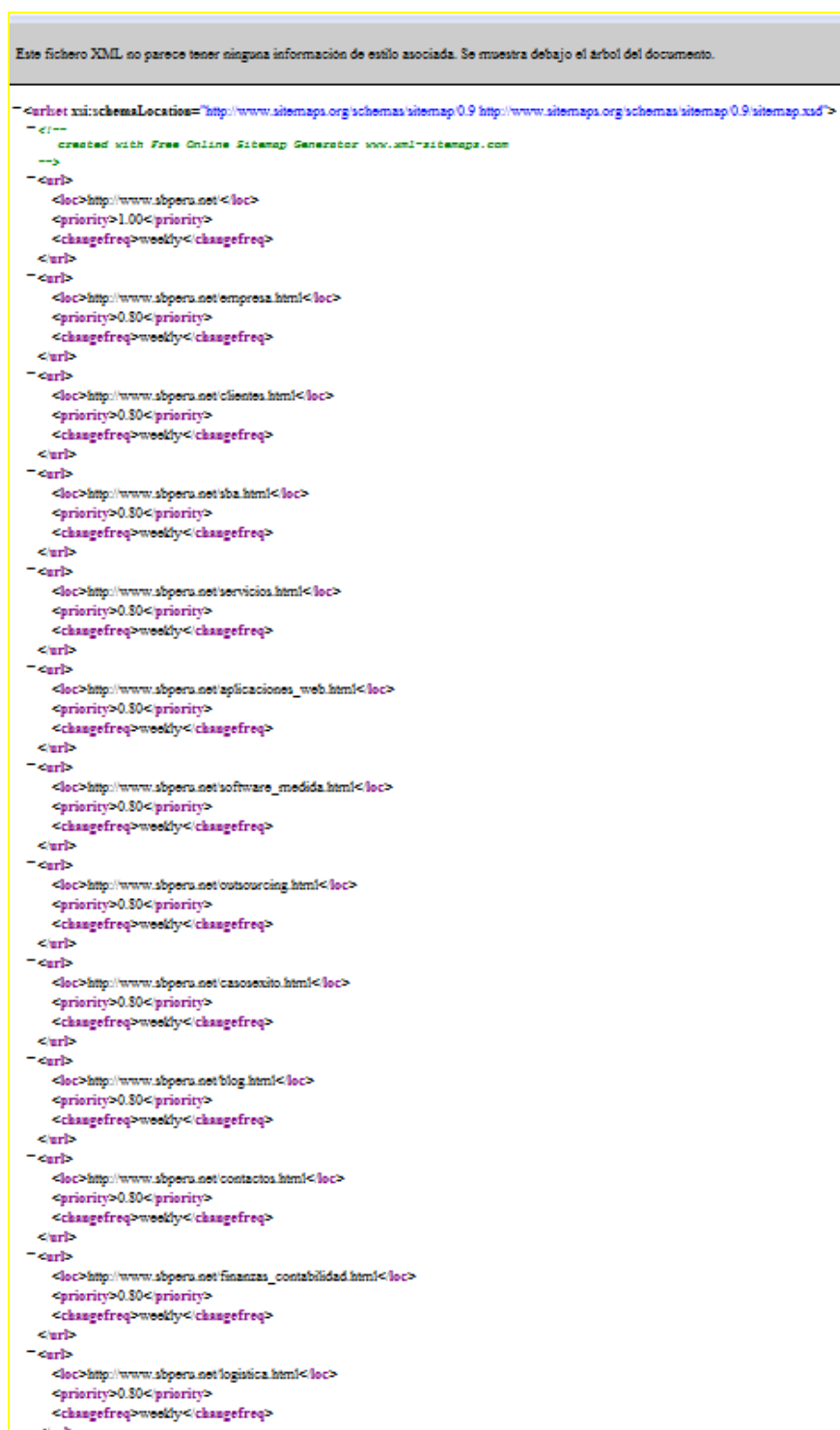


Figura 6.48: Archivo robots.txt para el portal de SoftBrilliance S.A.C.

● Archivo Sitemap.xml

Se creó el archivo sitemap.xml (véase la figura 6.49), con el cual hicimos que ciertas páginas Web, sean más fáciles de acceder por los motores de búsqueda, con la finalidad de que éstos puedan indexar rápidamente el contenido útil para usuarios navegantes.



```
<?xml version="1.0" encoding="UTF-8" ?>
<urlset xmlns:schemaLocation="http://www.sitemaps.org/schemas/sitemap/0.9 http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd"
  <!--
    created with Free Online Sitemap Generator www.xml-sitemaps.com
  -->
  <url>
    <loc>http://www.sbperu.net/</loc>
    <priority>1.00</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/empresa.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/clientes.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/sba.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/servicios.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/aplicaciones_web.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/software_medida.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/outourcing.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/casosexitos.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/blog.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/contactos.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/finanzas_contabilidad.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
  <url>
    <loc>http://www.sbperu.net/logistica.html</loc>
    <priority>0.80</priority>
    <changefreq>weekly</changefreq>
  </url>
</urlset>
```

Figura 6.49: Archivo sitemap.xml para el portal de SoftBrilliance S.A.C.

6.3.3 Análisis del Resultado de la Aplicación de la Solución al Caso de Estudio

Luego de haber reestructurado el Portal Web de SoftBrilliance S.A.C., se decidió realizar nuevamente un análisis y generar reportes de performance de palabras claves, que nos permitieran verificar la optimización del portal Web y comprobar los beneficios obtenidos.

A continuación mostramos los resultados de dicho análisis:

- Se comprobó que el portal Web sigue siendo indexado por los Motores de búsqueda

Lo primero a verificar es si las páginas del portal, continúan en un motor de búsqueda. Es decir, si aparece cuando alguien lo busca en Google, Yahoo y MSN.

Para esta prueba, se siguió lo siguientes pasos:

a. Examinar en la caché de los buscadores

Se buscó en Google, la página principal de SoftBrilliance S.A.C., luego se seleccionó la opción 'En Caché', tal como se muestra en la figura 6.50.



Figura 6.50: Búsqueda de la página principal del portal de SoftBrilliance en Google

En la figura 6.51, se ve que el buscador ha indexado el nuevo diseño de la página principal del portal de SoftBrilliance S. A. C.



Figura 6.51: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de Google

También, en la figura 6.52, se ve como es indexado en Yahoo y finalmente, en la figura 6.53, se ve como es indexado en MSN.

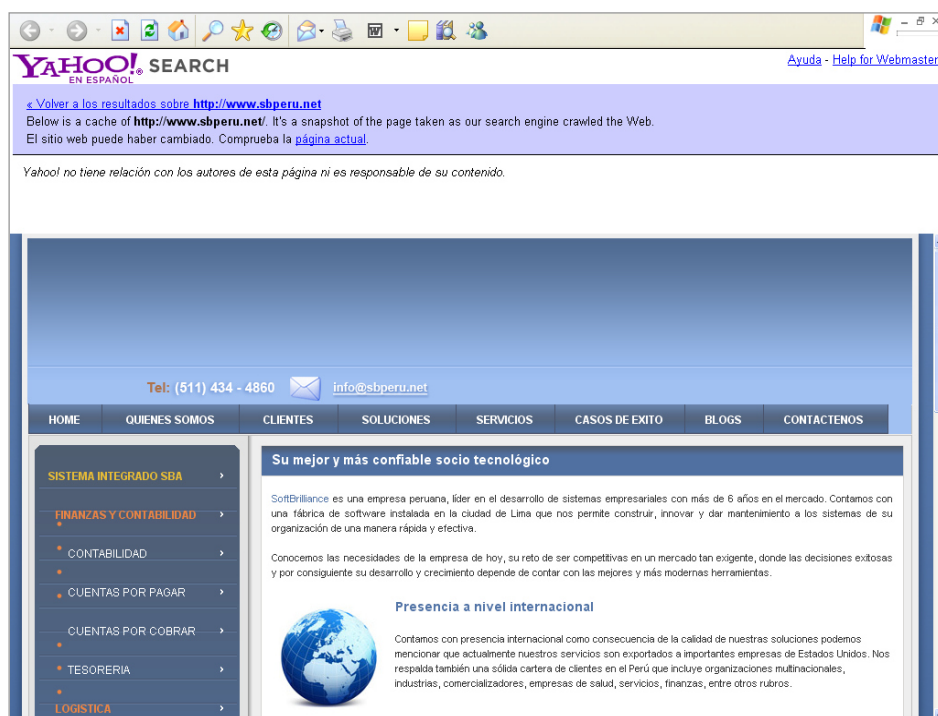


Figura 6.52: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de Yahoo



Figura 6.53: El nuevo diseño del Portal SofBrilliance S.A.C. ya está en la caché de MSN

En ese sentido, a diferencia del diseño anterior, este nuevo diseño es indexado y mostrado de una mejor manera, tanto en la cache de Google como la de Yahoo y MSN. Además se notó que a diferencia de Google, Yahoo, no tiene la capacidad de leer archivos flash, por lo cual queda justificada la atención que se debe presentar al contenido de las páginas, ver en la figura 6.52.

b. Ver el número de páginas indexadas

La siguiente prueba fue, obtener el número de páginas Web indexadas por los motores de búsqueda. Para lo cual se ingresó el texto 'site:www.sbperu.net', en la interfaz de consulta de cada uno de los buscadores.

Los resultados mostraron que Google, indexa un total de 165 páginas Web, tal como lo muestra la figura 6.54. Hubo un aumento del 60% del número de páginas Web indexadas.

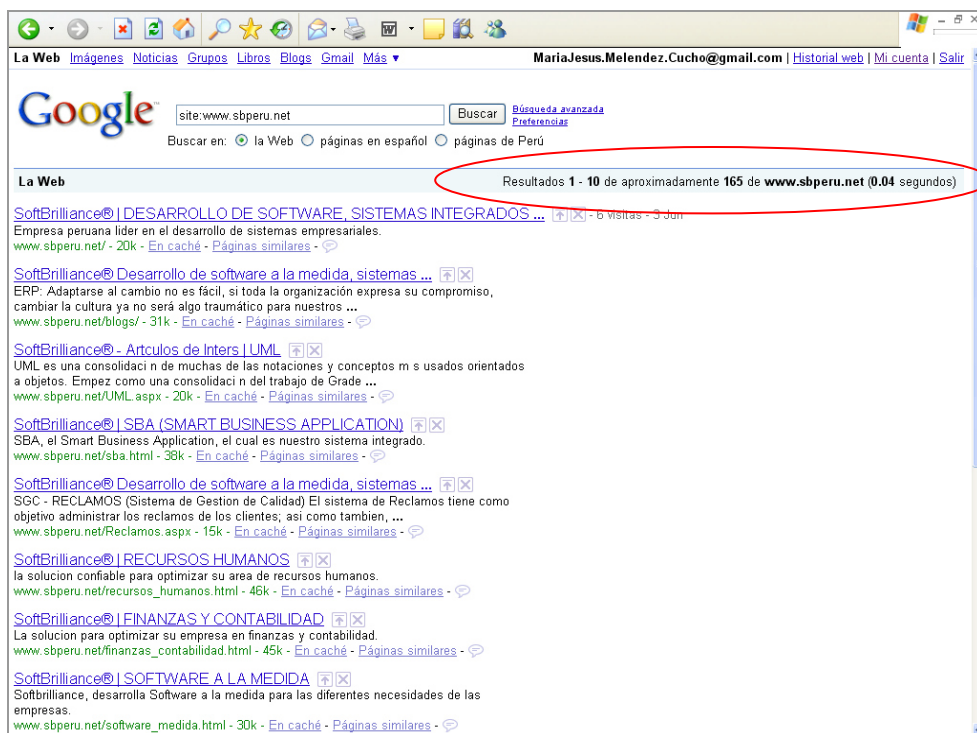


Figura 6.54: Número de páginas indexadas en Google.

Por otro lado, se tiene indexado 70 páginas Web en Yahoo, tal como se muestra en la figura 6.55.

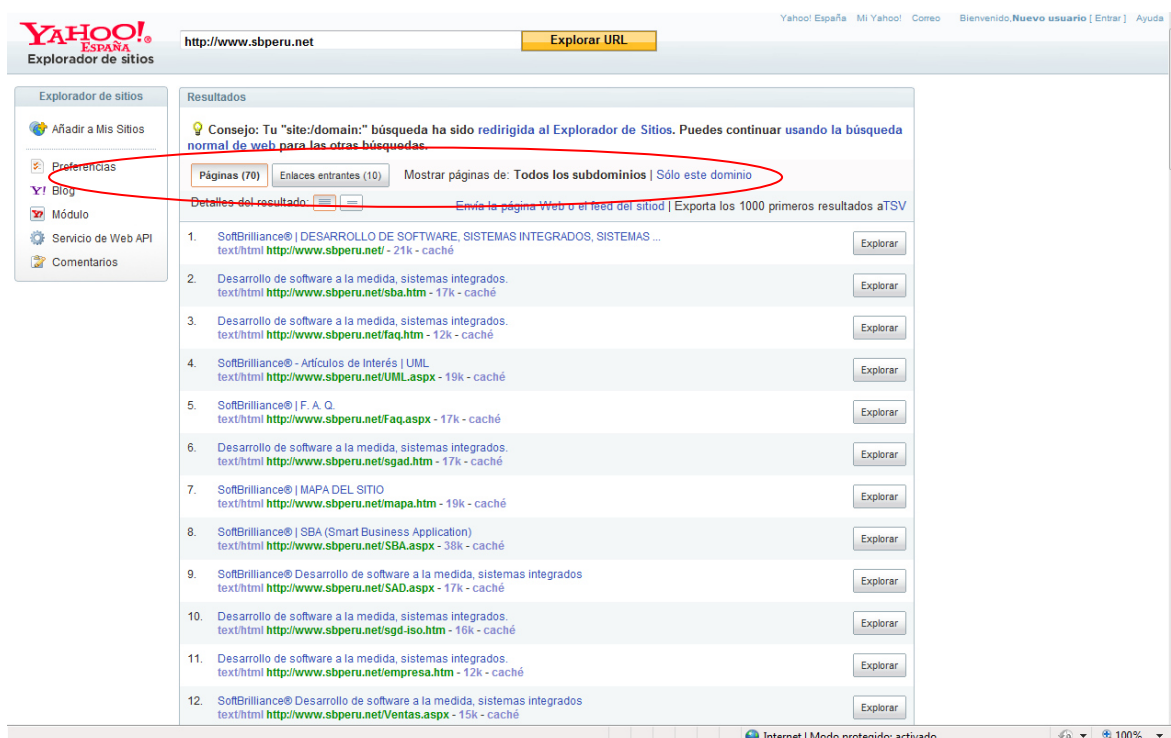


Figura 6.55: Número de páginas indexadas en Yahoo..

Finalmente y en menor número, se muestra 8 páginas Web indexadas por el motor de búsqueda, MSN, tal como se muestra en la figura 6.56.

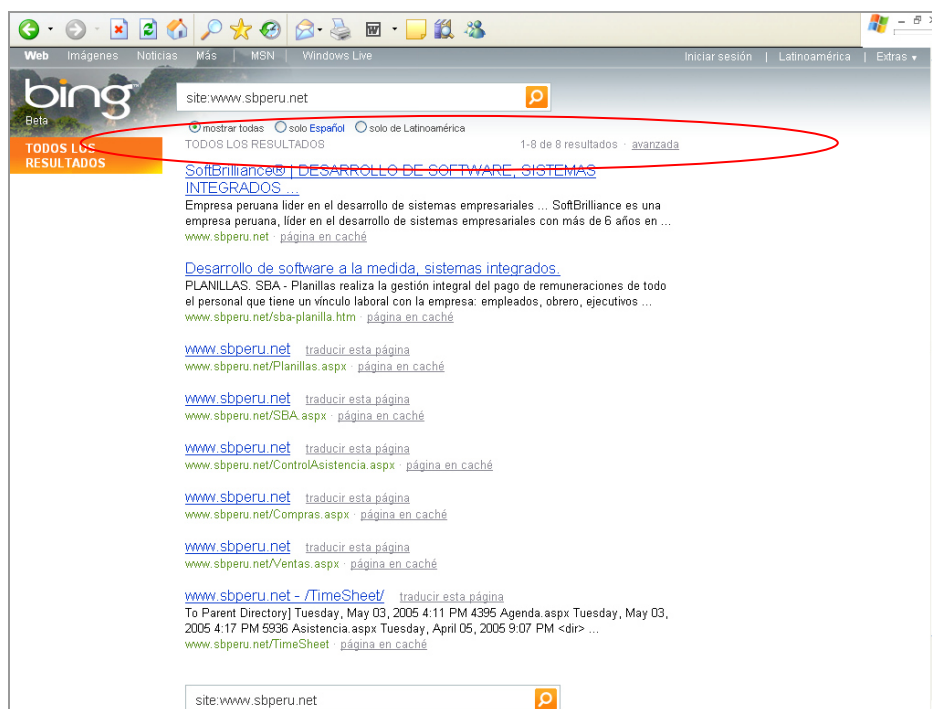


Figura 6.56: Número de páginas indexadas en MSN.

En conclusión, se puede afirmar que hubo un aumento en el número de páginas indexadas. En un análisis inicial del portal, se mostró que Google llegaba a indexar 56 páginas (véase la figura 6.26), pero actualmente indexa 165 páginas. Esto muestra un avance y una mejora del Portal Web, ya que el motor de búsqueda llegó a recorrer más páginas.

- **Se optimizó las páginas del portal**

En el anterior capítulo, se mencionó que la optimización del portal empezaría con el adecuado uso de los TAG TITTLE, METATAG DESCRIPTION etc...

Efectivamente, los motores de búsqueda empezaron a indexar las páginas Web, de acuerdo a los TAG TITTLE y METATAG DESCRIPTION, los que fueron ingresados por el sistema GECOSEO. En la figura 6.57, se muestra como se usó las propiedades para la pagina Web: 'mantenimiento.html'.

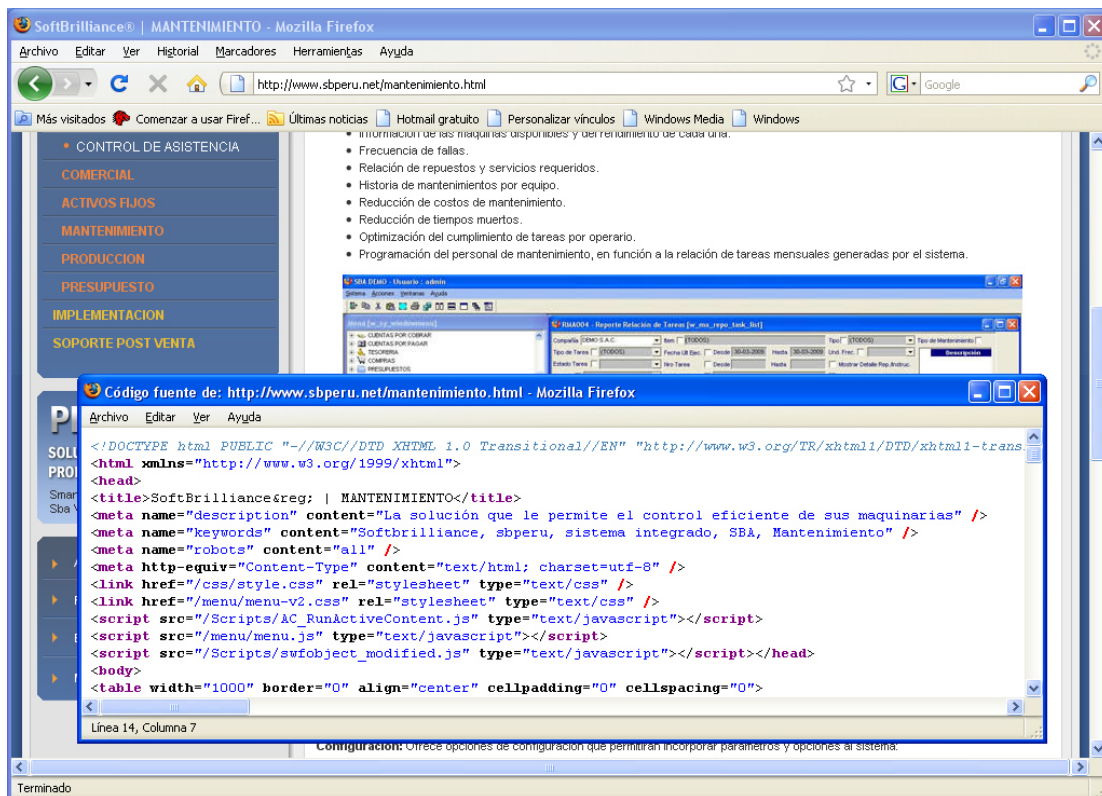


Figura 6.57: Propiedades de la Página Web “Mantenimiento.html” del Portal de SoftBrilliance S. A. C., configuradas por el sistema GECOSEO.

La indexación por los tags mencionados permitió el mejoramiento del posicionamiento del portal, ya que por medio de ellos, se da una visión general del contenido de las páginas Web indexadas en cada uno de los motores de búsqueda, tal como se visualiza en la figura 6.58.

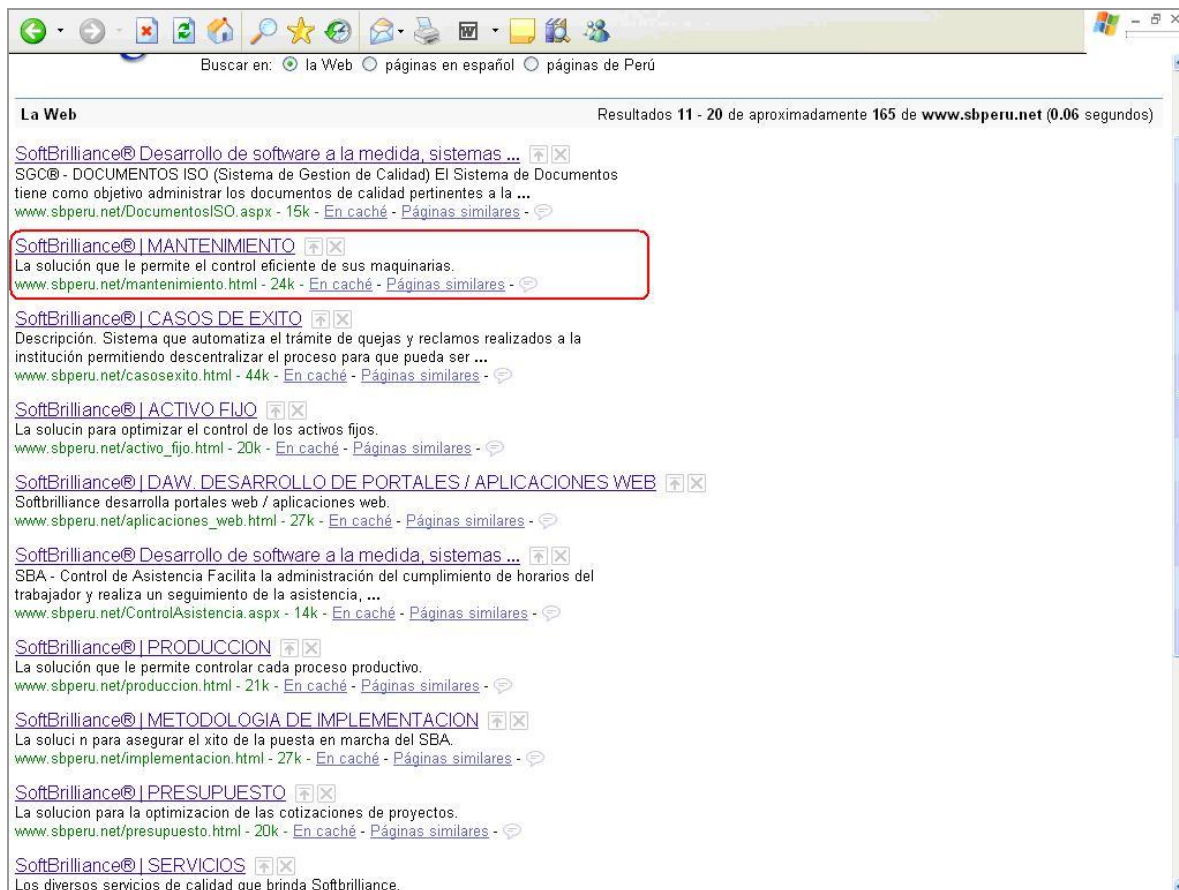


Figura 6.58: Optimización de la página Web 'Mantenimiento.html' mediante los TAG TITLE y METATAG DESCRIPTION.

- Se mejoró el posicionamiento en la que es mostrado el portal Web de SoftBrilliance

Con las actualizaciones y mejoras que se hizo al portal, se volvió a generar los reportes de performance en base a las palabras claves anteriormente seleccionadas (25 palabras claves) para los motores de búsqueda de Google, Yahoo y MSN respectivamente, tal como se muestra en las tablas 6.2, 6.3 y 6.4.

Keyword	GPos	GPag	GUrlFound	GPrevPos	GChange
aplicaciones inalambricas pdt	>100	>10		>100	Not Listed
consultoria informatica	>100	>10		>100	Not Listed
desarrollo de software	6	1	http://www.sbperu.net/	>100	Newly added
desarrollo de software a la medida	>100	>10		>100	Not Listed
integracion de sistemas	>100	>10		>100	Not Listed
outsourcing	>100	>10		>100	Not Listed
sba	>100	>10		>100	Not Listed
sbperu	1	1	http://www.sbperu.net/blogs/blogImages/	1	0
servicio de outsourcing	>100	>10		>100	Not Listed
servicios informaticos	>100	>10		>100	Not Listed
sistema de administracion de documentos	>100	>10		>100	Not Listed
sistema de gestion de documentos	>100	>10		>100	Not Listed
sistema de gestion de reclamos	28	3	http://www.sbperu.net/Reclamos.aspx	4	-24
sistema financiero contable	>100	>10		>100	Not Listed
sistema integrado	>100	>10		>100	Not Listed
sistemas a la medida	>100	>10		>100	Not Listed
sistemas empresariales	7	1	http://www.sbperu.net/	>100	Newly added
smart business application	>100	>10		>100	Not Listed
softbrilliance	1	1	http://www.sbperu.net/empresa.html	1	0
software a la medida	6	1	http://www.sbperu.net/	>100	Newly added
software development	>100	>10		>100	Not Listed
soluciones informaticas	>100	>10		>100	Not Listed
soluciones informaticas a la medida	>100	>10		>100	Not Listed
tecnologia informatica	>100	>10		>100	Not Listed
tecnologia informatica como herramienta estrategica	>100	>10		>100	Not Listed

Tabla 6.2: Reporte de Performance de Palabras Claves en Google.

Keyword	YPos	YPag	YUrlFound	YPrevPos	YChange
aplicaciones inalambbricas pdt	>100	>10		>100	Not Listed
consultoria informatica	>100	>10		>100	Not Listed
desarrollo de software	>100	>10		>100	Not Listed
desarrollo de software a la medida	7	1	http://www.sbperu.net/faq.htm	20	+13
integracion de sistemas	>100	>10		>100	Not Listed
outsourcing	>100	>10		>100	Not Listed
sba	42	5	http://www.sbperu.net/sba.htm	>100	Newly added
sbperu	1	1	http://www.sbperu.net/Ventas.aspx	1	0
servicio de outsourcing	>100	>10		>100	Not Listed
servicios informaticos	>100	>10		>100	Not Listed
sistema de administracion de documentos	26	3	http://www.sbperu.net/sba.htm	>100	Newly added
sistema de gestion de documentos	50	5	http://www.sbperu.net/sba.htm	>100	Newly added
sistema de gestion de reclamos	6	1	http://www.sbperu.net/sba.htm	6	0
sistema financiero contable	46	5	http://www.sbperu.net/sba.htm	>100	Newly added
sistema integrado	>100	>10		>100	Not Listed
sistemas a la medida	9	1	http://www.sbperu.net/faq.htm	>100	Newly added
sistemas empresariales	7	1	http://www.sbperu.net/	>100	Newly added
smart business application	13	2	http://www.sbperu.net/sba.htm	>100	Newly added
softbrilliance	1	1	http://www.sbperu.net/	10	+9
software a la medida	1	1	http://www.sbperu.net/faq.htm	>100	Newly added
software development	>100	>10		>100	Not Listed
soluciones informaticas	>100	>10		>100	Not Listed
soluciones informaticas a la medida	>100	>10		>100	Not Listed
tecnologia informatica	>100	>10		>100	Not Listed
tecnologia informatica como herramienta estrategica	>100	>10		>100	Not Listed

Tabla 6.3: Reporte de Performance de Palabras Claves en Yahoo.

Keyword	MPos	MPag	MUrlFound	MPrevPos	MChange
aplicaciones inalamblicas pdt	>100	>10		>100	Not Listed
consultoria informatica	>100	>10		>100	Not Listed
desarrollo de software	>100	>10		>100	Not Listed
desarrollo de software a la medida	45	5	http://www.sbperu.net/	18	-27
integracion de sistemas	>100	>10		>100	Not Listed
outsourcing	>100	>10		>100	Not Listed
sba	>100	>10		>100	Not Listed
sbperu	1	1	http://www.sbperu.net/TimeSheet	1	0
servicio de outsourcing	>100	>10		>100	Not Listed
servicios informaticos	>100	>10		>100	Not Listed
sistema de administracion de documentos	>100	>10		>100	Not Listed
sistema de gestion de documentos	>100	>10		>100	Not Listed
sistema de gestion de reclamos	>100	>10		65	Dropped out
sistema financiero contable	>100	>10		>100	Not Listed
sistema integrado	>100	>10		>100	Not Listed
sistemas a la medida	>100	>10		37	Dropped out
sistemas empresariales	>100	>10		82	Dropped out
smart business application	>100	>10		>100	Not Listed
softbrilliance	6	1	http://www.sbperu.net/	9	+3
software a la medida	>100	>10		31	Dropped out
software development	>100	>10		>100	Not Listed
soluciones informaticas	>100	>10		>100	Not Listed
soluciones informaticas a la medida	>100	>10		>100	Not Listed
tecnologia informatica	>100	>10		>100	Not Listed
tecnologia informatica como herramienta estrategica	>100	>10		>100	Not Listed

Tabla 6.4: Reporte de Performance de Palabras Claves en MSN.

En donde cada estado de la columna 'XChange' significa lo siguiente:

Newly added	Nuevamente añadido la palabra clave a las 100 primeras posiciones.
+XXX	Subió XXX posiciones con respecto a su posición anterior.
0	Sigue en la misma posición que su posición anterior.
-XXX	Bajó XXX pociones con respecto a su posición anterior.
Dropped out	Salió la palabra clave fuera de las 100 primeras posiciones

A continuación, se muestran gráficos que explicarán mejor, el nuevo posicionamiento del portal de Softbrilliance S.A.C después de su optimización. Este posicionamiento es en Google, Yahoo y MSN. Cabe mencionar que el análisis más completo, se encuentra adjuntó en el Anexo C.

a. Posicionamiento en Google

Los nuevos reportes mostraron que Google posiciona al portal de SoftBrilliance en la primera página, con un 16 % de palabras claves, tal como se muestra en la figura 6.59.

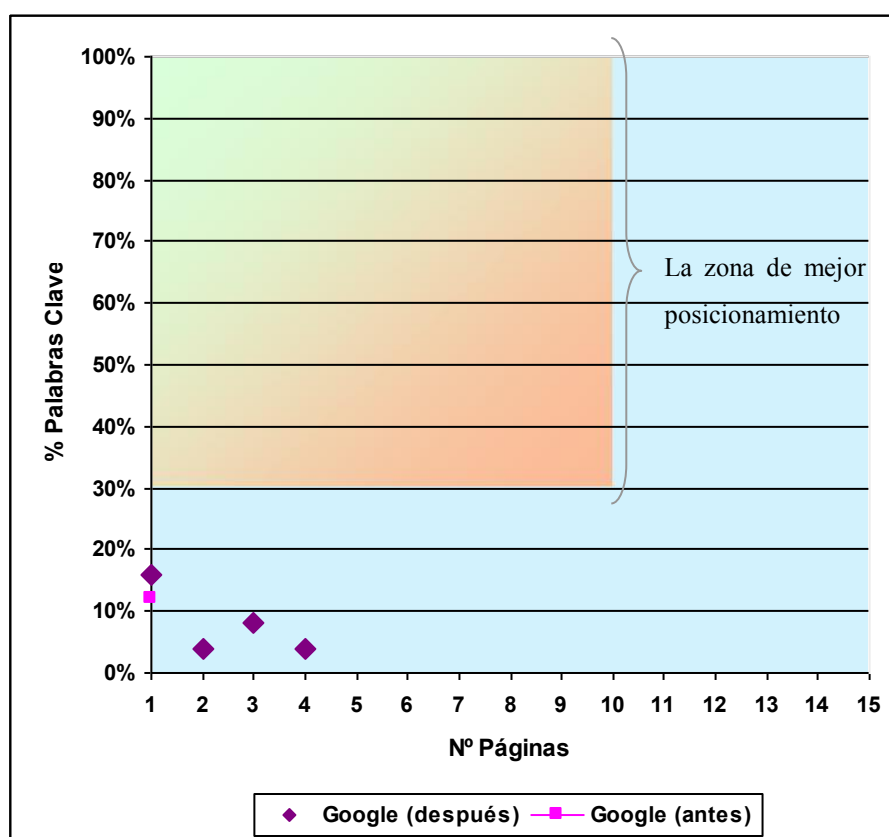


Figura 6.59: Puntos de posicionamiento del portal Web en Google, luego de la aplicación de la Técnica WCMS y SEO.

Por otro lado, en la figura 6.60, se muestra la proporción de palabras claves, mediante las cuales, el portal es posicionado.

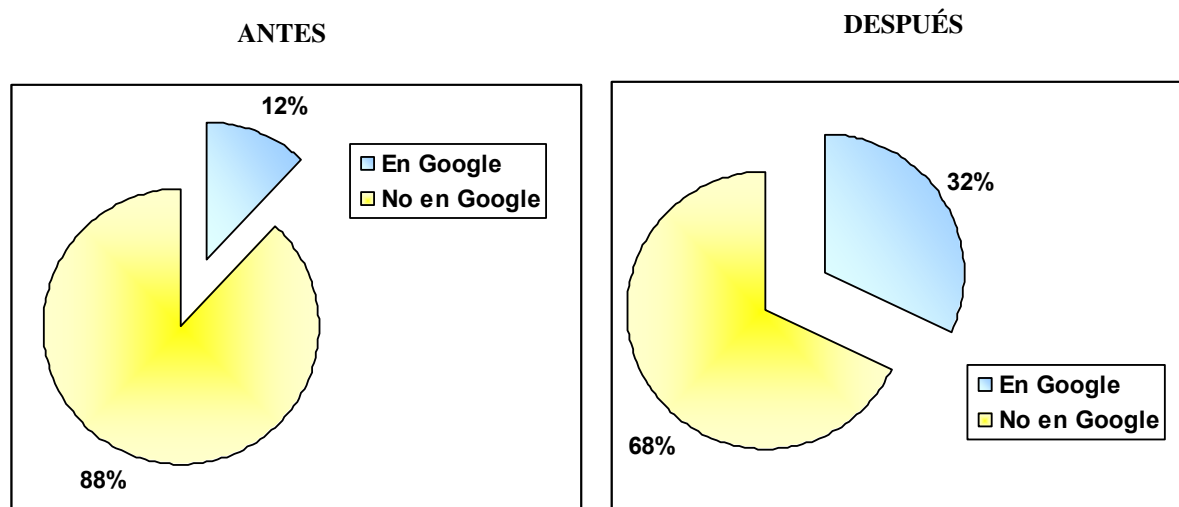


Figura 6.60: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en Google, luego de la aplicación de la Técnica WCMS y SEO.

En ese sentido, podemos concluir, que el posicionamiento en Google se está mejorando, yaque hubo un aumento del 20%, en el número de palabras claves por la cual es indexado el portal.

b. Posicionamiento en Yahoo

Al igual que con Google, se generó un nuevo reporte a fin de mostrar como Yahoo está posicionando al portal de SoftBrilliance. El resultado mostró que no sólo se logró aumentar de un 16% a un 48 % las palabras claves por la cual es posicionado, sino se logró subir en el posicionamiento, es decir es indexado en más páginas de las SERPs, tal como se muestra en la figura 6.61.

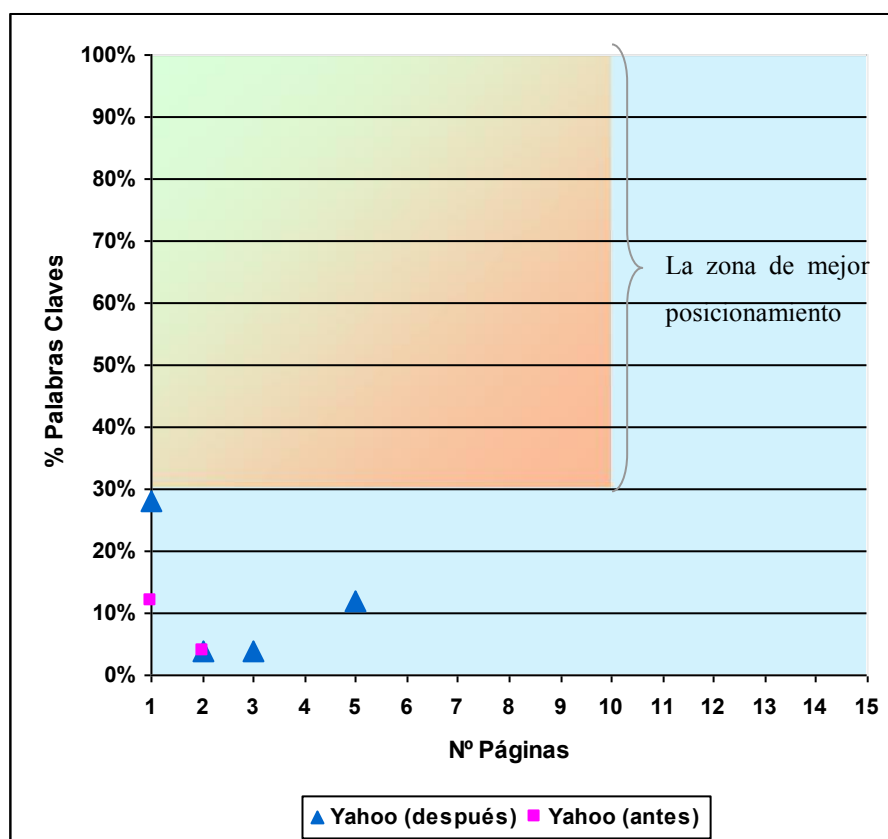


Figura 6.61: Punto de posicionamiento del portal Web en Yahoo, luego de la aplicación de la Técnica WCMS y SEO.

Además, en la figura 6.62, se puede observar la anterior y la actual proporción de palabras claves, por el cual, el portal es posicionado.

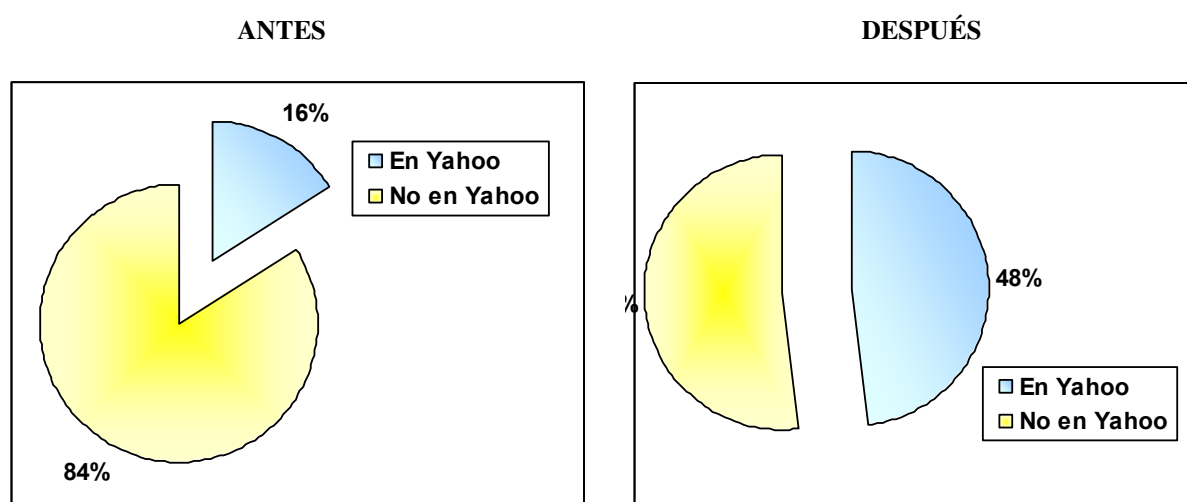


Figura 6.62: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en Yahoo, luego de la aplicación de la Técnica WCMS y SEO.

En ese sentido, podemos concluir, que el posicionamiento en Yahoo, ha mejorado notablemente, ya que hubo un aumento del 32%, en el número de palabras claves por la cual es indexado el portal y estamos llegando a la zona de mejor posicionamiento.

c. Posicionamiento en MSN

Al igual que los otros Motores de búsqueda, se generó el reporte de cómo MSN llega a posicionar al portal de SoftBrilliance, como se muestra en la figura 6.63.

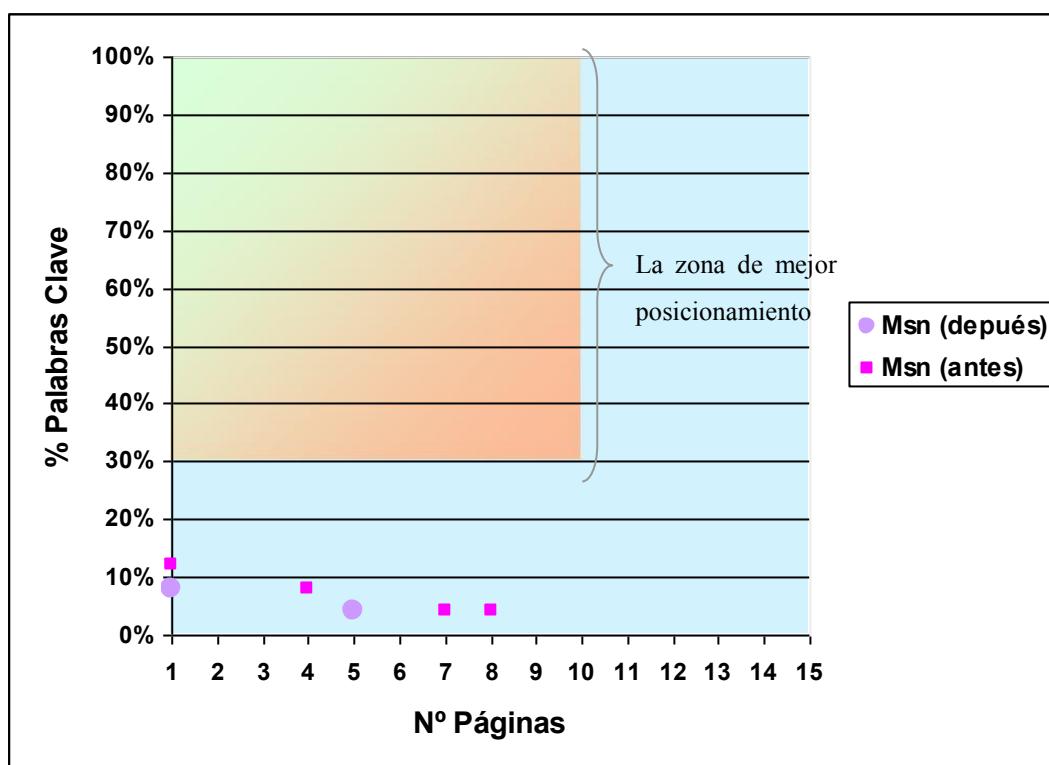


Figura 6.63: Punto de posicionamiento del portal Web en MSN, luego de la aplicación de la Técnica WCMS y SEO.

Debido a que MSN, ha cambiado a un nuevo motor de búsqueda, llamado “Bing” (<http://www.bing.com>), su robot (MSNBot), aún no recorre por completo los sitios Web de la Internet. Esto se ve reflejado en la disminución del número de palabras claves por la cual es indexado el portal.

Este escenario mejorará en el transcurso de las semanas e incluso meses, cuando el MSNBot termine de recorrer el portal Web.

En la figura 6.64 se puede observar la proporción de palabras claves, por el cual, el portal es posicionado para este motor de búsqueda.

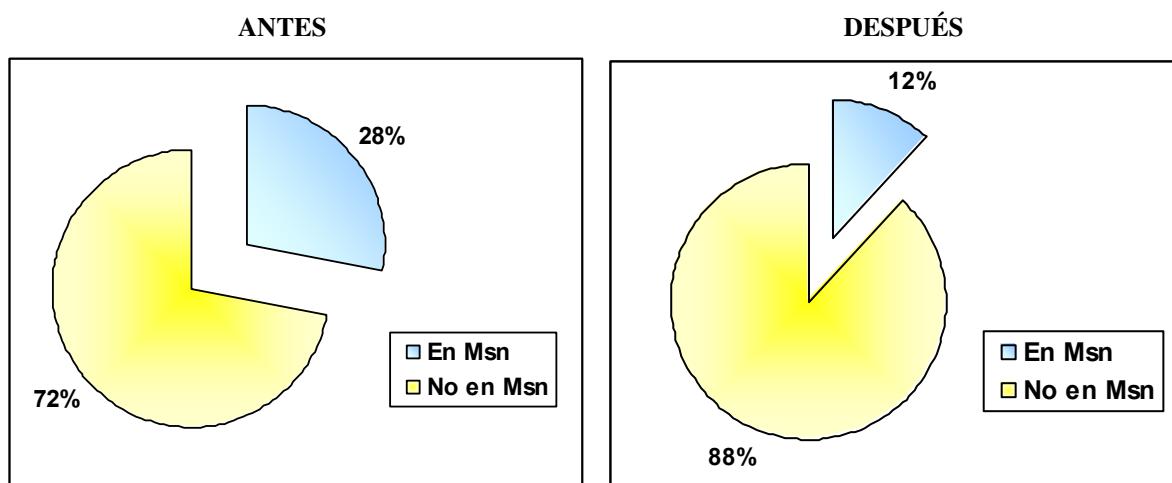


Figura 6.64: Comparación del anterior y el actual porcentaje de palabras claves por las que es posicionado en MSN, luego de la aplicación de la Técnica WCMS y SEO

- **Se comprobó la mejora en el posicionamiento, mediante los Reportes Estadísticos de Google Analytics**

Para un mejor análisis y seguimiento del posicionamiento Web del portal de SoftBrilliance S. A. C., se generó una cuenta en Google Analytics (ver figura 6.65); que es un servicio de estadísticas, el cual te permite recopilar, monitorear, ver y analizar datos sobre el tráfico de tu sitio Web; tales como: el número de visitas, regiones de donde provienen las búsquedas, el porcentaje de tráfico en tu portal y otras estadísticas.

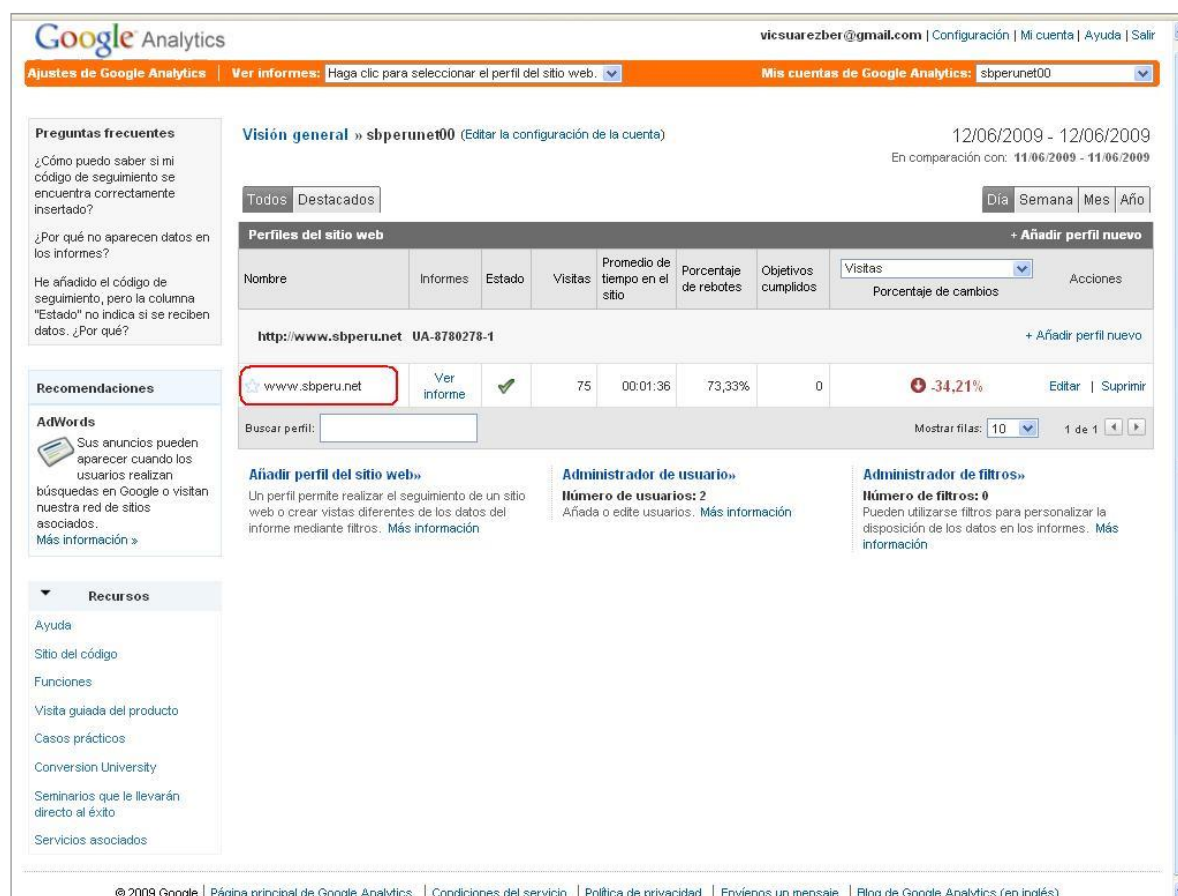


Figura 6.65: Cuenta en Google Analytics del Portal del Portal de SoftBrilliance S. A. C.

A continuación se muestran algunos reportes estadísticos de Google Analytics, para el portal de SoftBrilliance S.A.C.

Estos reportes estadísticos corresponde a 5 semanas (10/05/2009 – 10/06/2009), luego de la aplicación de la propuesta de solución. Estos datos ayudarán a determinar futuras acciones ya sea para mejorar, modificar o corregir las estrategias de posicionamiento.

En la figura 6.66, se muestra como los usuarios visitantes han estado haciendo uso del portal, durante este tiempo de validación.

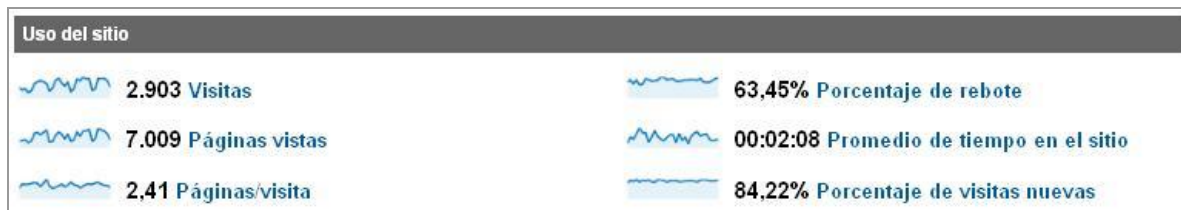


Figura 6.66: Cuenta en Google Analytics del Portal del Portal de SoftBrilliance S. A. C.

En donde:

- *Visitas*: representa el número de sesiones individuales iniciadas por todos los usuarios para llegar al portal de SoftBrilliance.
- *Páginas Vistas*: representa el número de veces que un navegador carga una página del portal de SoftBrilliance.
- *Páginas / Visitas*: viene a ser el promedio de páginas, que los usuarios visualizan durante una única sesión, en el portal de SoftBrilliance.
- *Promedio de Tiempo en el Sitio*: representa el tiempo de permanencia en una página del portal de SoftBrilliance.
- *Porcentaje de Visitas Nuevas*: representa el porcentaje de usuarios nuevos que visitan el portal de SoftBrilliance.

Adicionalmente, en el reporte estadístico de la figura 6.67, podemos visualizar un incremento por semana del número de visitantes del portal de SoftBrilliance.

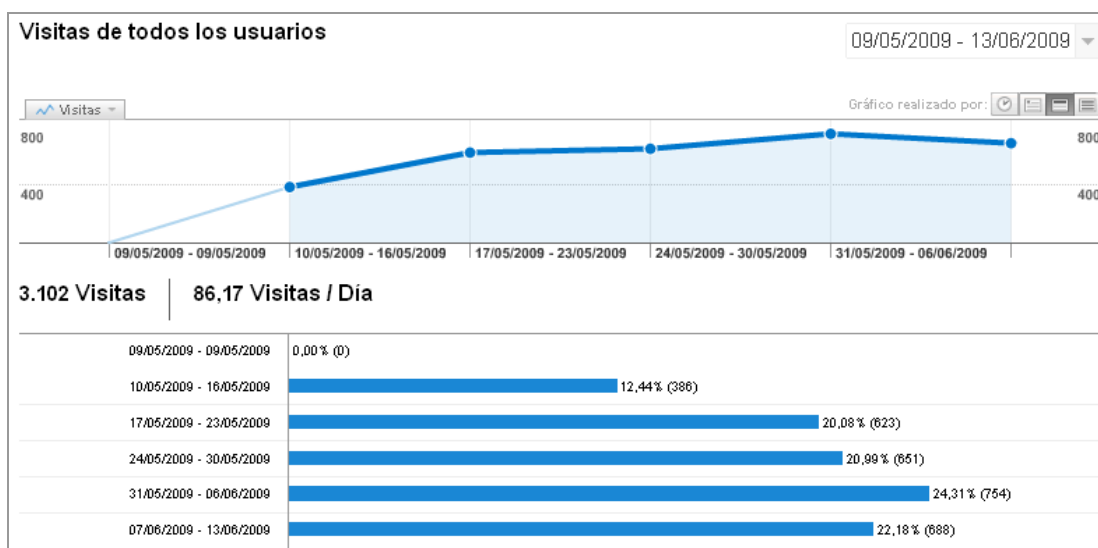


Figura 6.67: Reporte Estadístico del Número de Visitas Semanalmente del Portal de SoftBrilliance S. A. C.

Así también, se visualiza un incremento de páginas vistas semanalmente, tal como se muestra en la figura 6.68.

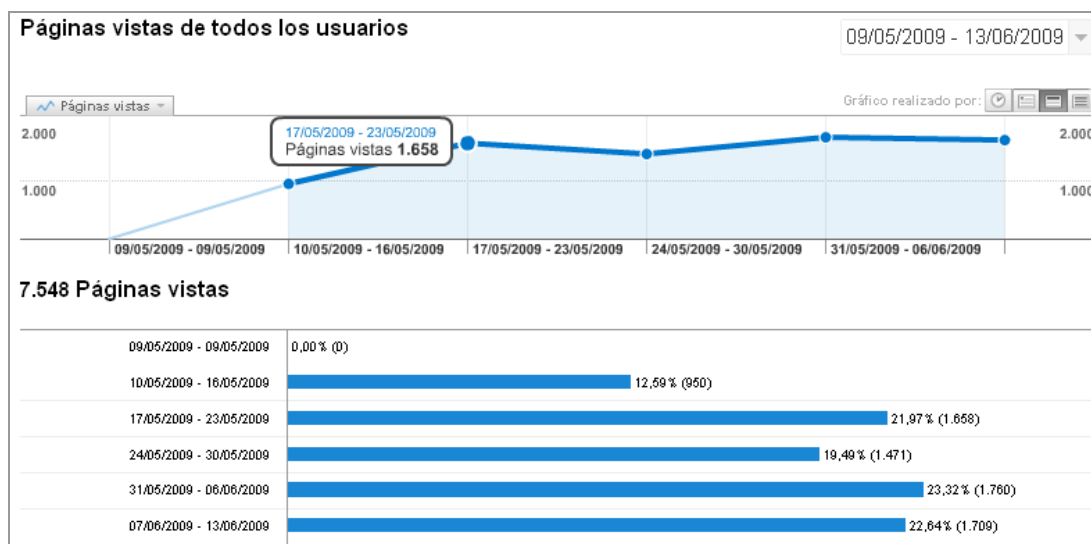


Figura 6.68: Reporte Estadístico del Número de Visitas Semanalmente del Portal de SoftBrilliance S. A. C.

Y para complementar estos 2 reportes anteriores de Google Analytics, podemos visualizar, en la figura 6.69, un reporte en el cual se muestra el tráfico de visitas por día de la semana en el portal de SoftBrilliance.

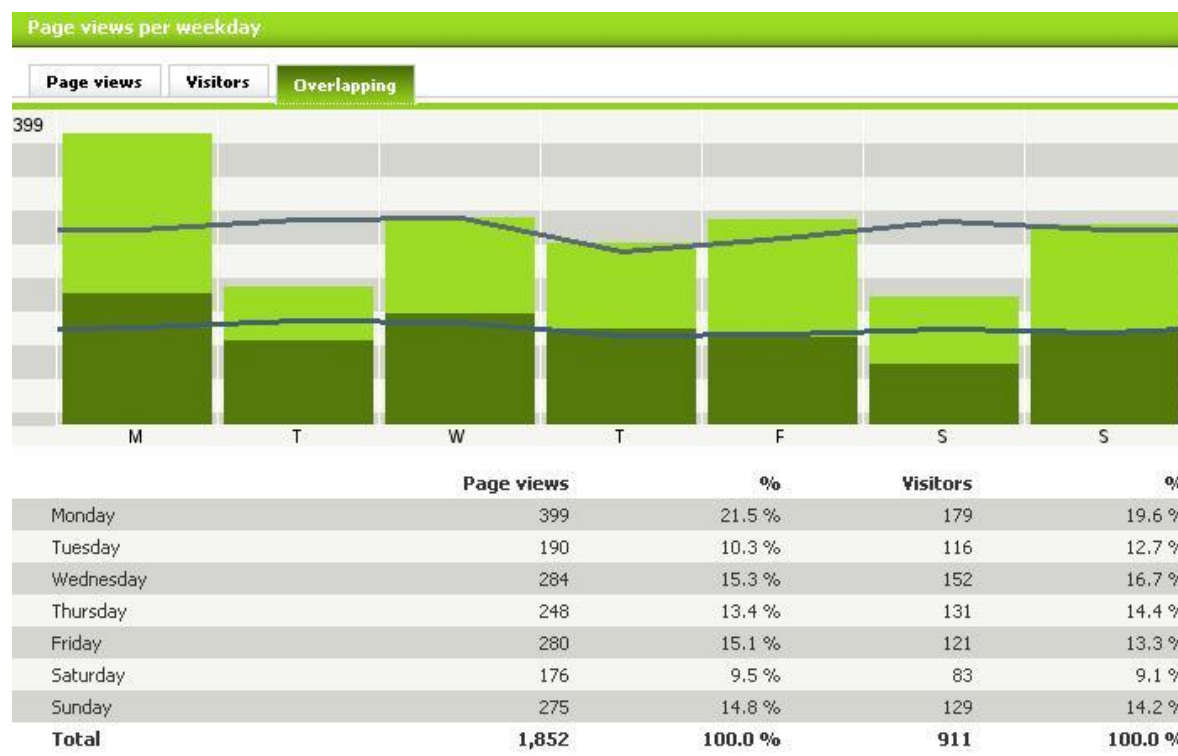


Figura 6.69: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.

Por otro lado, en cuanto a las fuentes de tráfico hacia el portal de SoftBrilliance, conseguimos un 70% de tráfico por los motores de búsqueda, tal como se muestra en la figura 6.70.



Figura 6.70: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.

Y este tráfico por parte de los motores de búsqueda se ha conseguido por 7 buscadores, obteniendo un mayor porcentaje del motor de búsqueda, Google (97%), así como se muestra en la figura 6.71.

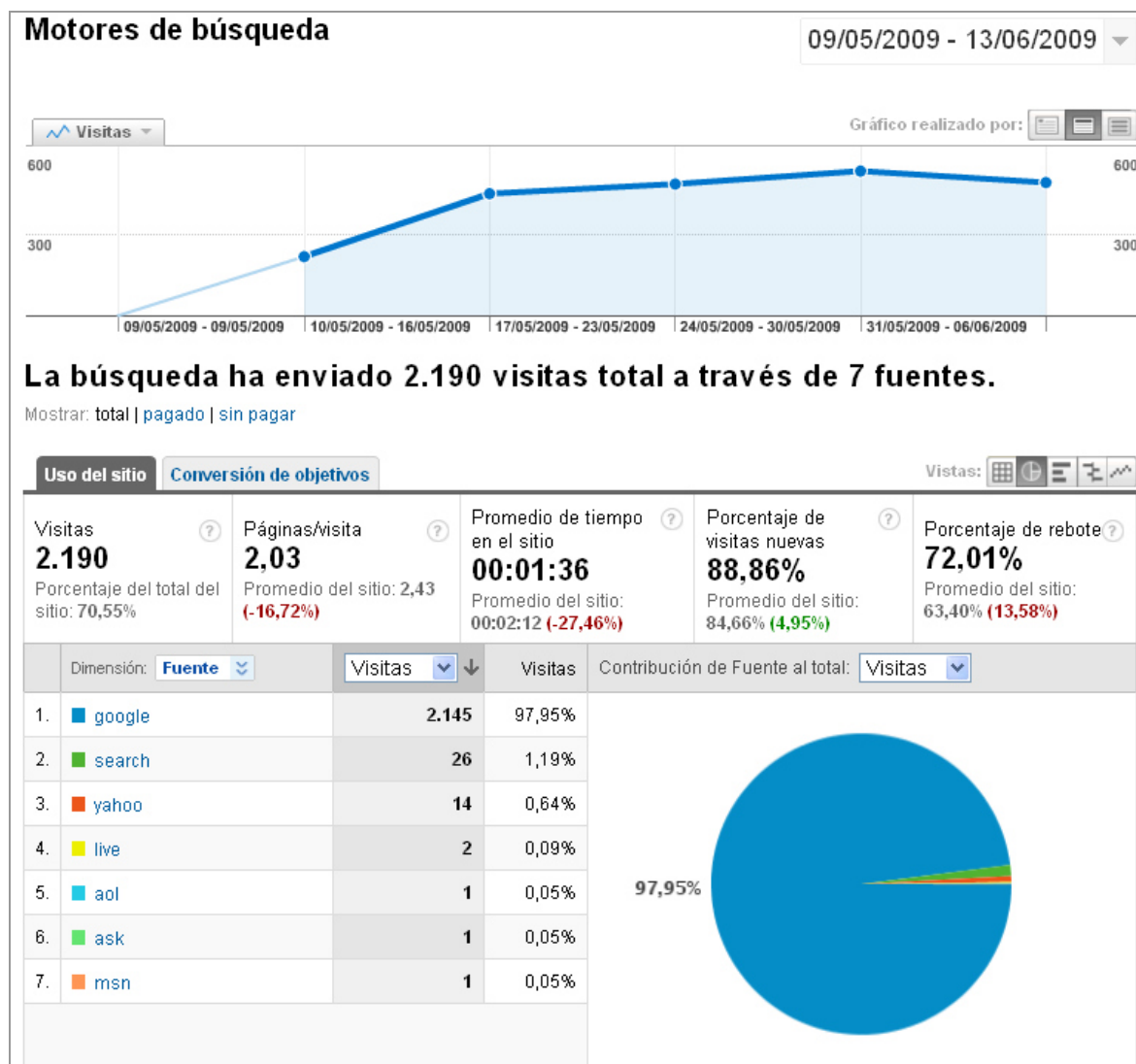


Figura 6.71: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.

Gráfico de visitas por ubicación

Averigüe la procedencia de los usuarios que visitan su sitio Web e identifique los mercados geográficos más lucrativos.

Orientación geográfica

3.120 visitas provinieron de 31 países/territorios.

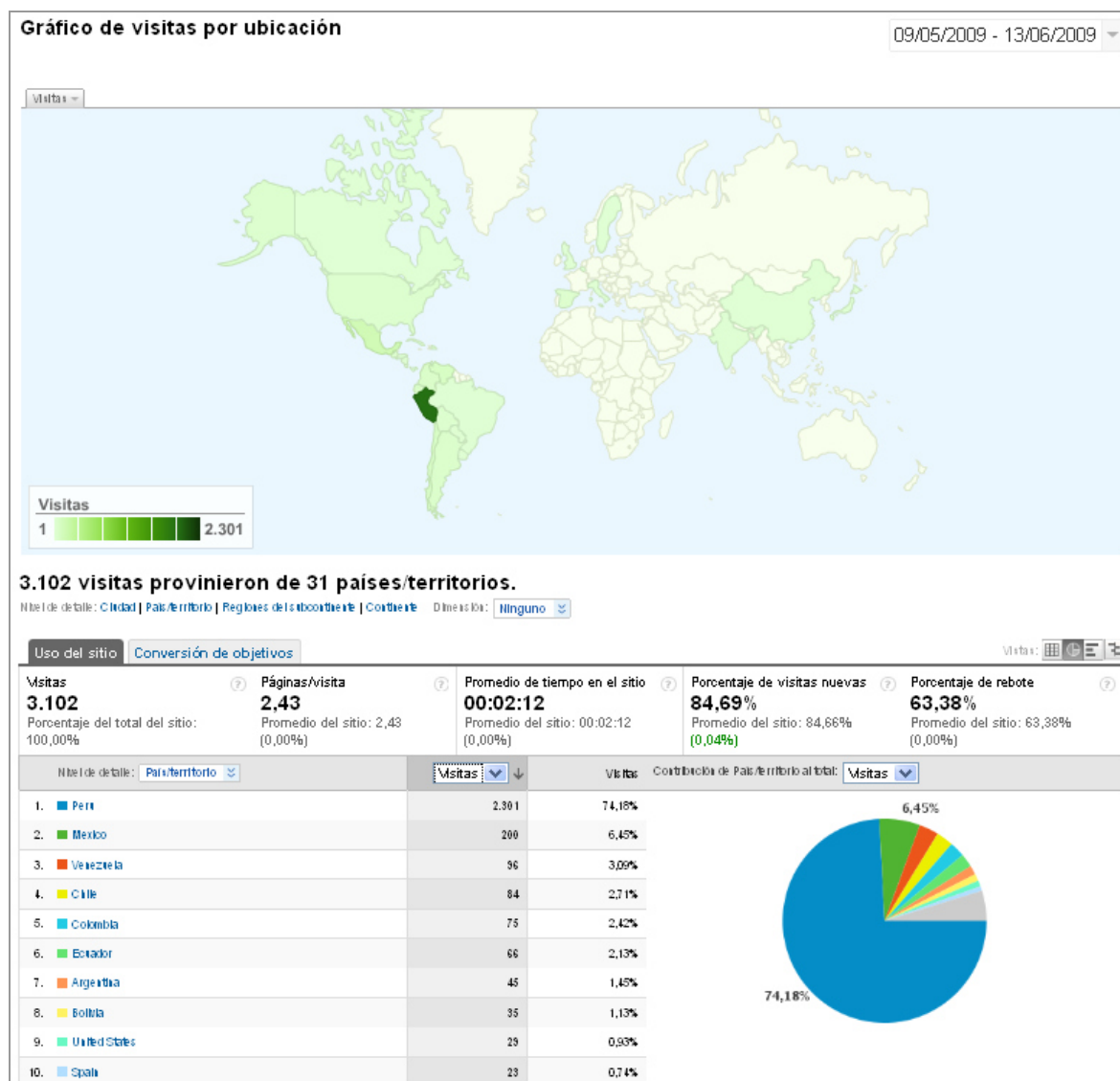


Figura 6.72: Gráfico de Visitas por Ubicación Geográfica del Portal de SoftBrilliance S. A. C.

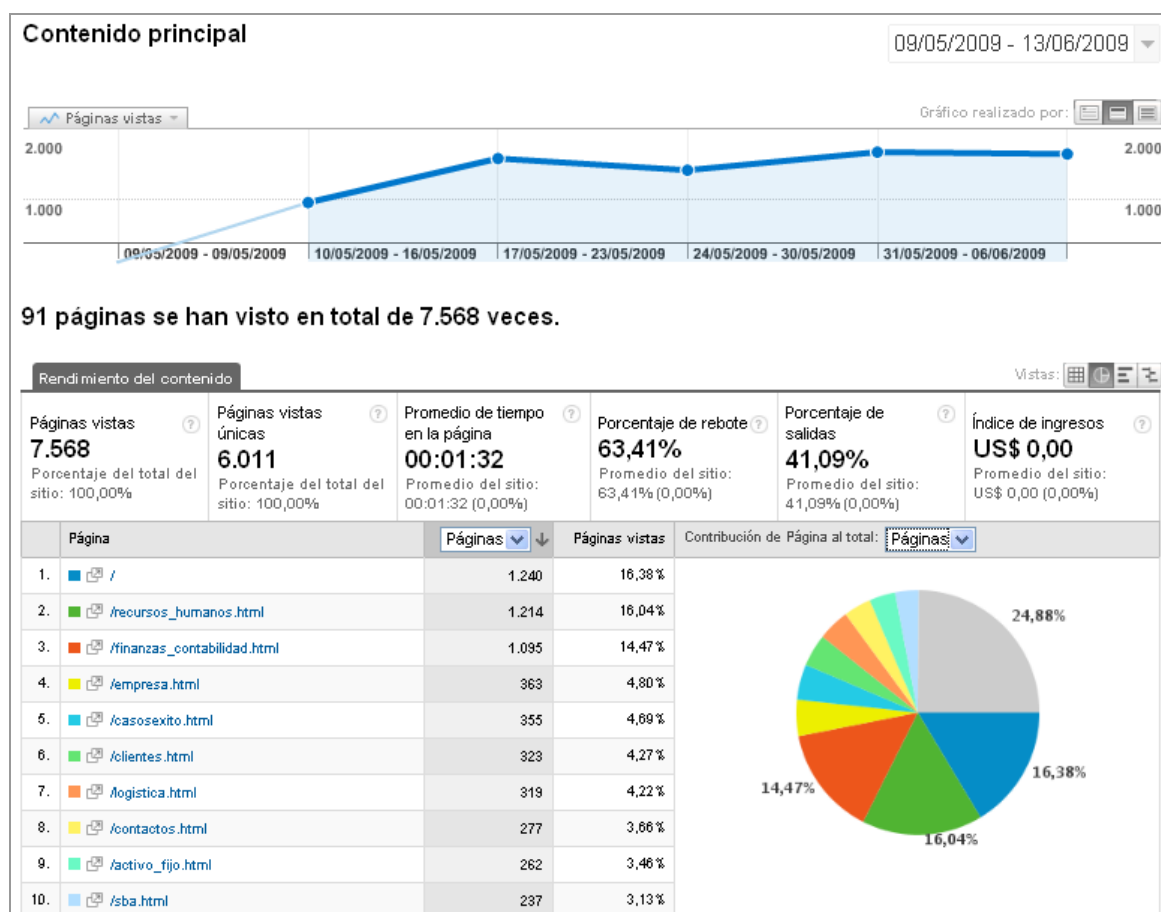


Figura 6.73: Estadísticas del Tráfico de Visitas por Día de la Semana del Portal de SoftBrilliance S. A. C.

- **Resumen del resultado de la aplicación**

Por todo lo visto anteriormente, el análisis de las páginas y los reportes de los motores de búsqueda, se puede llegar a las siguientes conclusiones:

- Se logró un aumento de las páginas indexadas, esto significa una mejora debido a que los motores de Búsqueda dispondrán de un mayor contenido que mostrar en cada búsqueda.
- Se ha mejorado el posicionamiento Web del portal de SoftBrilliance S. A. C, en solo 5 semanas. Pero debido a que los Motores de Búsqueda muestran resultados en un promedio de 6 meses se puede prever que estos resultados irán mejorando, pues así lo demuestra el aumento importante en el posicionamiento que tuvo el portal en las primeras 5 semanas.
- Se ha conseguido optimizar el uso de palabras claves, así lo demuestra el hecho que los Motores de Búsqueda indexan a partir de los TAG TITTLE Y LOS METATAGs, los cuales contienen dichas palabras claves.
- Los Motores de Búsqueda logran indexar con un mayor número de palabras claves.
- Se ha conseguido mayor número de usuarios visitantes al portal Web.
- Se está logrando que el portal Web no sólo sea amigable para el usuario visitante (atractivo), sino también a los motores de búsqueda.

6.4 Resumen

Se construyó un prototipo de acuerdo a la propuesta definida por este trabajo de investigación, el cual consistía en la integración de dos técnicas que permiten optimizar el diseño y la información tanto para los Motores de Búsqueda como para el usuario.

Dentro de las funcionalidades del prototipo figuran principalmente, la creación de páginas Web en base a las normas del buen diseño y a las recomendaciones SEO, además la de gestionar las propiedades y elementos del portal, y finalmente permite gestionar, de forma uniforme y dinámica, la información mostrada en el portal.

Posteriormente se eligió el portal Web de una empresa peruana para poder aplicar el prototipo, como primer punto se analizó el estado actual del portal, es decir, ver si las páginas estaban optimizadas para los motores de Búsqueda. Además se realizó reportes que demostraron que el portal no estaba adecuadamente optimizado para los Motores de Búsqueda.

Eso se debía en parte al poco conocimiento del tema que tenía la empresa y a que fue diseñado para el usuario, es decir el portal era amigable sólo al usuario más no era amigable a los Motores de Búsqueda.

Tras este análisis se aplicó el prototipo propuesto al portal Web, se modificó el Contenido Web de cierta manera que la información esté optimizada y actualizada, de igual manera y mediante el prototipo se ingresó los TAGs con las palabras claves correspondientes en cada página Web, es decir se optimizó la estructura y sus propiedades.

Se agregó los archivos sitemap.xml y el robot.txt que permitirían guiar a los robots de los motores de búsqueda recorrer de mejor manera el portal.

Finalmente, se hizo pruebas finales que permitirían demostrar la optimización del portal, así como la mejora en la optimización en los motores de Búsqueda. No sólo se logró el aumento de las páginas indexadas, sino la mejora del posicionamiento Web del portal de SoftBrilliance S. A. C. Se puede decir que el portal fue optimizado tanto para los Motores De Búsqueda como para los usuarios.

Capítulo 7: Conclusiones y Trabajos Futuros

En esta sección, se dará a conocer las conclusiones a las cuales se llegó luego de la realización del presente trabajo de investigación. Además, se dará a conocer aquellos temas que puedan ser tomados en siguientes investigaciones.

7.1 Conclusiones

- Se ha demostrado que las anteriores propuestas no llegan a optimizar por completo los portales Web, es decir muchas de ellas se orientan exclusivamente a optimizar para los Motores de Búsqueda dejando de lado gestión de Contenidos y otros optimizan la información del portal pero olvidando los Motores de Búsqueda.
- La nueva propuesta logró optimizar el Portal Web siguiendo las recomendaciones SEO y gestionando el contenido. Esta optimización fue completa, es decir, fue tanto para los motores de búsqueda como para los usuarios.
- La optimización completa del Portal Web, logró una mejora en el posicionamiento Web en los principales Motores de Búsqueda.
- Se ha demostrado que la Gestión de Contenidos también forma parte de la optimización para los Motores de Búsqueda, debido a que les proporciona más información de donde escoger.
- Se demostró que los buscadores no sólo son simples herramientas de búsqueda sino un medio más para las estrategias de marketing.
- Existe un aumento en la necesidad de acercar las empresas al tema de posicionamiento de una forma sencilla, que les permita obtener un mayor provecho en comparación con sus competidores más cercanos.
- SEO puede llegar a ser un punto de apoyo a las técnicas actualmente existentes. En base a éstas, se puede definir nuevas formas de aplicación de técnicas para la optimización de las páginas de un portal Web.

- La optimización de la información en el portal Web colabora con los motores de búsqueda, por lo cual se concluye que cada uno busca la optimización del portal, por lo que resultan ser complementarias.
- El posicionamiento es, por decirlo de alguna manera, una ciencia experimental, donde los grandes Motores de Búsqueda protegen sus algoritmos de gestión de información, obligando a los portales Web a evolucionar en su posicionamiento mediante procesos de prueba y error.
- Los contenidos de calidad bien estructurados y relacionados son el elemento clave del posicionamiento.
- El posicionamiento depende y dependerá de la forma en la que los buscadores organicen la información, en esto radica la importancia de estar actualizado, tanto en la información que se maneja como en las nuevas técnicas o tendencias, como nuevos algoritmos de búsqueda, nuevos buscadores, etc.

7.2 Trabajos Futuros

Existen posibles trabajos o investigaciones que se pueden continuar

- **SEO para eCommerce**

Los Motores de Búsqueda no indexan muy frecuentemente los sitios de comercio electrónico. Esto se debe a que los portales Web que ofrecen sus productos en línea, ofrecen poco contenido útil (existe un mayor manejo de imágenes), la mayoría no incluye descripciones de productos. Debido a todo esto, se puede llegar a concluir que los sitios de comercio electrónico no son amigables a los motores de búsqueda.

Es así que, para siguientes investigaciones, queda permitir a sitios Web dedicados al comercio electrónico, poder ser optimizado para los motores de búsqueda, así como mejorar la gestión de contenidos e información.

- **Mobile SEO**

Debido al continuo desarrollo de páginas Web para móviles, los diseñadores de páginas Web no conocen por completo todo lo referente al diseño y estándares en las páginas Web para móviles. Además, los motores de búsqueda en los móviles son nuevas y en proceso de desarrollo. Por estas razones, móviles SEO y SEM son, aún experimentales.

Pero aún así los objetivos en los buscadores en móviles son los mismos. Se desea tráfico, y aún más que los usuarios puedan encontrarlos en los motores de búsqueda. Por lo anterior, Mobile SEO viene a ser un tema para ser tomado en las siguientes investigaciones.

Referencias Bibliográficas

- [1] Arbildi Larreina, Iñigo. *Posicionamiento en buscadores: una metodología práctica de optimización de sitios web*. El profesional de la información. Vol. 14, N° 2, (Marzo – Abril, 2005), pp. 108 – 124.
- [2] Wendy Chisholm, Gregg Vanderheiden, Ian Jacobs, Web Content Accessibility Guidelines 1.0, Consorcio World Wide Web (W3C), <http://www.w3.org/TR/WCAG10/>, [consulta: 20 de Enero del 2009].
- [3] Beatriz Aguilera Escamilla, *Sistemas Avanzados de Recuperación de la Información*, Universidad Carlos III, Trabajo, (2006), Madrid - España.
- [4] CTPCBA – Colegio de Traductores Públicos de la Ciudad de Buenos Aires, *Motores de búsqueda*, Coleccionables de Aplicaciones Informáticas, Edición 5. (2006 – 2008). pp. 1 – 3.
- [5] Cristòfol Rovira, Lluís Codina, Mari-Carmen Marcos, Rafael Pedraza, *Máster Online en Buscadores. Selección de unidades didácticas 2007/2008*, Creative Commons del tipo: Reconocimiento – No comercial – Sin obras derivadas, (septiembre, 2008), España.
- [6] Francisco Javier Martínez Méndez, *Propuesta y desarrollo de un modelo para la evaluación de la recuperación de información en internet*, Universidad de Murcia, Tesis de Doctorado, (2002), Murcia - España.
- [7] Chang, G. Norwell, *Mining the WorldWide Web: an information search approach*, Kluwer Academic publishers, (2001), Massachussets - Estados Unidos.
- [8] Aguilar González, R, *Monografía sobre motores de búsqueda*, Yahoo Geocities, <http://www.geocities.com/motoresdebusqueda/inicio.html>, [consulta: 22 de Febrero del 2009].
- [9] Adelaida M. Delgado Domínguez, *Mecanismos de Recuperación de Información en la WWW*, Universidad de las Islas Baleares, (Junio, 1998), Palma - España.

- [10] M. Koster, *Robots in the Web: threat or treat*, <http://info.webcrawler.com/mak/projects/robots/threat-or-treat.html>, (1995), [consulta: 30 de Enero del 2009].
- [11] D. Sullivan, *Search Engine Features Chart*, <http://www.searchenginewatch.com/features.htm>, (1997), [consulta: 20 de Diciembre del 2008].
- [12] Gil Leiva, I, *La automatización de la indización de documentos*, Gijón: Trea, (1999), pp. 221.
- [13] Sergey Brin y Lawrence Page, *The anatomy of a large-scale hypertextual web search engine*, Computer Networks and ISDN Systems, Vol. 30, (1998), pp. 107 – 117.
- [14] Arauzo Galindo, M. *Nuevas técnicas de búsqueda en Internet*. Universidad Complutense, http://www.fdi.ucm.es/asignaturas/ssii_sup/SI/grupo1/resumenes/Nuevas%20tecnicas%20de%20busqueda.htm, (2000), [Consulta: 10 de diciembre de 2008].
- [15] Codina, Lluís y Marcos, Mari Carmen. *Posicionamiento web: conceptos y herramientas*. El profesional de la información, Vol 14, Nº 2, (marzo - abril, 2005), pp. 84.
- [16] David de Ugarte, *El libro del posicionamiento en buscadores*, (2002), pp. 5.
- [17] Codina, Lluís y Marcos, Mari Carmen. *Posicionamiento web: conceptos y herramientas*. El profesional de la información, Vol 14, Nº 2, (marzo - abril, 2005), pp. 85.
- [18] Peter Kent. *Search Engine Optimizarion for Dummies*. Wiley Publishing, Inc., (2004), Indiana – Estados Unidos.
- [19] Pablo Lara Navarra, José Ángel Martínez Usero, *Agentes inteligentes en la búsqueda y recuperación de información*, Planeta – UOC, (Julio, 2006), Barcelona - España.
- [20] Jennifer Grappone y Gradiva Couzin, *Search Engine Optimization An Hour a Day*, Wiley Publishing, Inc., (2008), New Jersey - Estados Unidos.

- [21] Kristopher B. Jones, *Search Engine Optimization: Your visual blueprint™ for effective Internet marketing*, Wiley Publishing, Inc., (2008), Indiana - Estados Unidos.
- [22] López Gómex, Miguel, *Marketing Online: Posicionamiento en Buscadores*, Creative Commons, publicable, copiable, distribuible de cualquier modo, No editable, No comercial, (2008, 2009), Valencia - España.

Índice de Anexos

- [1] **Anexo A:** Documento de Especificación de Casos de Uso del Sistema.
- [2] **Anexo B:** Análisis del Posicionamiento Web Inicial del portal de SoftBrilliance S.A.C.
- [3] **Anexo C:** Análisis del Posicionamiento Web, luego de la Aplicación de la Solución.
- [4] **Anexo D:** Reportes Estadísticos de Google Analytics para el portal de SoftBrilliance S.A.C.